

THESIS / THÈSE

MASTER EN SCIENCES MATHÉMATIQUES

L'algorithme de biorthogonalisation de Lanczos et ses beakdowns

DUFOUR, Cécile

Award date:
1996

Awarding institution:
Universite de Namur

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Facultés universitaires Notre-Dame de la Paix
Namur
Facultés des Sciences :
Département de Mathématiques

**L'algorithme de biorthogonalisation
de Lanczos
et ses "breakdowns".**

Mémoire présenté pour l'obtention du grade
de Licencié en Sciences
mathématiques
par

Promoteur : Mr J-P Thiran

Cécile Dufour

Année Académique 1995-1996

Que toutes les personnes qui m'ont soutenues pendant ces années d'étude trouvent ici l'expression de ma gratitude.

Je remercie Monsieur le Professeur Jean-Pierre Thiran, promoteur de ce mémoire, pour l'aide et les conseils qu'il m'a prodigués lors de la réalisation de ce mémoire.

Je tiens également à remercier mon père Pierre Dufour pour la patience et l'aide précieuse dont il a témoigné au cours de ce travail.

Enfin, merci aux élèves de seconde licence pour la bonne ambiance qui a régné durant ces quatre années d'étude. Merci à toute ma famille et mes amis pour leur soutien et leurs encouragements.

L'algorithme de biorthogonalisation de Lanczos et ses breakdowns.

Nom : Cécile Dufour.

Promoteur : Mr J-P Thiran.

Résumé :

Partant de deux vecteurs initiaux et d'une matrice carrée A , l'algorithme de biorthogonalisation de Lanczos est un processus permettant d'engendrer deux séquences finies de vecteurs biorthogonaux. Il peut être utilisé notamment pour la résolution des systèmes d'équations linéaires. Cet algorithme peut donner lieu à un phénomène appelé "breakdown" provoquant l'arrêt prématuré du processus. Deux types de "breakdown" existent : l'un qualifié de "sérieux" et l'autre d'"arrêt prématuré".

Ce mémoire proposera une manière d'éviter l'apparition de tels phénomènes par le biais de l'étude des propriétés d'orthogonalité et les formules de récurrence des polynômes orthogonaux ainsi que par la théorie des approximants de Padé.

The Lanczos biorthogonalization algorithm and his breakdowns.

Abstract :

Beginning with two initial vectors and a $N \times N$ matrix A , the Lanczos biorthogonalization algorithm is a process for computing two finite sequences of biorthogonal vectors. This process can notably be used for the resolution of linear systems of equations. This algorithm can give rise to a phenomenon called "breakdown", inducing the premature stopping of the process. Two types of breakdowns exist : the first is called "serious breakdown" and the other "premature termination".

This paper will propose a way to avoid the occurrence of such phenomenons by means of the study of orthogonality and recurrence formulas of formal orthogonal polynomials as well as by means of the Padé approximants theory.

Table des matières

Introduction	5
1 L'algorithme de biorthogonalisation de Lanczos.	7
1.1 Préliminaires	7
1.2 L'algorithme de biorthogonalisation (BO).	10
1.3 Les différents "breakdowns".	11
1.3.1 Arrêt prématuré.	12
1.3.2 "Breakdown" sérieux.	14
1.3.3 Remarque.	14
1.4 Algorithme de résolution de systèmes d'équations linéaires.	14
2 Les polynômes orthogonaux formels et les approximants de Padé.	17
3 Les formules d'orthogonalité et de récurrence des polynômes orthogonaux formels.	31
4 Interprétation matricielle.	40
5 Les polynômes orthogonaux et l'algorithme de biorthogonalisation de Lanczos.	46

5.1	L'orthogonalité et l'algorithme de Lanczos	46
5.2	Les formules de récurrence et l'algorithme de Lanczos	48
5.3	Algorithme non générique de biorthogonalisation.	50
6	Les "breakdowns" de l'algorithme de biorthogonalisation de Lanczos.	52
6.1	Formulation matricielle des chapitres 1 et 3.	52
6.2	Les "breakdowns" et la table de Padé.	54
6.3	L'algorithme non générique de biorthogonalisation de Lanczos pour la résolution des systèmes linéaires	59
	Conclusion	62
	Bibliographie	63

Introduction

Le processus de Lanczos appliqué à une matrice carrée A est un algorithme permettant d'engendrer deux séquences finies de vecteurs mutuellement orthogonaux. Il permet notamment de résoudre des systèmes d'équations linéaires.

Lors de l'application de cet algorithme, deux phénomènes appelés "breakdown" peuvent survenir. Ceux-ci provoqueront l'arrêt prématuré du processus. L'un des "breakdowns" sera qualifié de "sérieux" et l'autre, de "prématuré".

Dans ce mémoire, nous analysons une manière d'éviter de tels phénomènes. Pour ce faire, nous nous baserons sur deux articles de Gutknecht [4] et [5]. L'étude des propriétés d'orthogonalité et des formules de récurrence des polynômes orthogonaux ainsi que la théorie des approximants de Padé [2] nous serviront d'intermédiaires pour formuler l'algorithme remédiant à l'apparition de "breakdowns".

Voici en quelques mots les différentes étapes du mémoire.

Dans le premier chapitre, nous formulerons l'algorithme de biorthogonalisation de Lanczos dans le cas générique, c'est-à-dire, en l'absence de "breakdown" [4]. Nous proposerons également un algorithme pour la résolution des systèmes d'équations linéaires.

Le second chapitre étudiera en parallèle les polynômes orthogonaux et la théorie des approximants de Padé [2].

Dans le troisième chapitre, nous nous pencherons sur l'étude des propriétés d'orthogonalité des polynômes et des formules de récurrence qui s'en déduisent [5].

Le chapitre quatre interprétera, sous forme matricielle, les différents résultats obtenus

dans les deux chapitres précédents.

Le chapitre cinq est une reformulation de l'algorithme générique par l'introduction de la théorie des éléments étudiés dans les chapitres précédents.

Pour terminer, le chapitre six sera consacré à l'étude des "breakdowns" de l'algorithme non générique de biorthogonalisation de Lanczos, présenté au chapitre précédent.

Chapitre 1

L'algorithme de biorthogonalisation de Lanczos.

1.1 Préliminaires

Soit $A \in \mathbb{C}^{N \times N}$, une matrice complexe d'ordre N et $x_0, y_0 \in \mathbb{C}^N$, un couple de vecteurs initiaux qui ne sont pas orthogonaux, i.e.,

$$\langle y_0, x_0 \rangle \neq 0$$

où $\langle y_0, x_0 \rangle$ est le produit scalaire Euclidien $\tilde{y}_0 x_0$ avec \tilde{y}_0 , le transposé conjugué de y_0 . L'algorithme de biorthogonalisation de Lanczos [7] consiste à déterminer deux suites finies de vecteurs $x_n (0 \leq n \leq \nu - 1)$ et $y_n (0 \leq n \leq \nu - 1)$ dans des sous-espaces de Krylov de A et \tilde{A} respectivement,

$$\begin{aligned} x_n &\in K_n = \text{span}\{x_0, Ax_0, \dots, A^n x_0\}, \\ y_n &\in L_n = \text{span}\{y_0, \tilde{A}y_0, \dots, \tilde{A}^n y_0\} \end{aligned} \quad (1.1)$$

qui sont biorthogonaux.

$$\langle y_m, x_n \rangle = \begin{cases} 0 & \text{si } m \neq n \\ \delta_n \neq 0 & \text{si } m = n. \end{cases} \quad (1.2)$$

La longueur ν dépend de A , x_0 et y_0 . Elle résulte de l'impossibilité de prolonger les suites de sorte que (1.1), (1.2) soient satisfaits avec $\delta_n \neq 0$: on parlera d'une rupture ou "breakdown" du processus.

Nous avons la propriété suivante:

Propriété 1.1

$$K_n = \text{span}\{x_0, x_1, \dots, x_n\}, \quad L_n = \text{span}\{y_0, y_1, \dots, y_n\}.$$

Preuve. Nous démontrerons, par récurrence, la propriété pour K_n , le raisonnement étant identique pour L_n . La propriété est évidente pour $n = 0$. Si elle est vraie pour $n - 1$, $K_{n-1} = \text{span}\{x_0, x_1, \dots, x_{n-1}\}$. Nous obtenons $K_n = \text{span}\{x_0, \dots, x_{n-1}, A^n x_0\}$. Le vecteur $x_n \in K_n$ s'écrit

$$x_n = \sum_{j=0}^{n-1} \alpha_j x_j + \alpha_n A^n x_0$$

où $\alpha_n \neq 0$ sinon y_n est orthogonal à x_n et nous serions en contradiction avec (1.2). On déduit donc

$$A^n x_0 = \alpha_n^{-1} (x_n - \sum_{j=0}^{n-1} \alpha_j x_j)$$

de sorte que $K_n = \text{span}\{x_0, \dots, x_{n-1}, x_n\}$. □

On déduit immédiatement pour $0 < n < \nu$,

$$x_n \perp L_{n-1}, \quad y_n \perp K_{n-1} \quad (1.3)$$

et

$$x_n \in K_n \setminus K_{n-1}, \quad y_n \in L_n \setminus L_{n-1}. \quad (1.4)$$

Soit $x_{n-1} \in K_{n-1} \setminus K_{n-2}$, nous pouvons écrire $x_{n-1} = \sum_{j=0}^{n-1} \alpha_j A^j x_0$ où $\alpha_{n-1} \neq 0$. Le vecteur Ax_{n-1} où $n < \nu$ s'écrit

$$Ax_{n-1} = \sum_{j=1}^n \alpha_{j-1} A^j x_0$$

de sorte que $Ax_{n-1} \in K_n \setminus K_{n-1}$. Par conséquent,

$$Ax_{n-1} = \sum_{j=0}^n \tau_{j,n-1} x_j, \quad \tau_{n,n-1} \neq 0, \quad (1.5)$$

et, semblablement,

$$\tilde{A}y_{n-1} = \sum_{j=0}^n \sigma_{j,n-1} y_j, \quad \sigma_{n,n-1} \neq 0. \quad (1.6)$$

Pour $n = \nu$, (1.5) et (1.6) peuvent être satisfaits pour $x_\nu \perp L_{\nu-1}$ et $y_\nu \perp K_{\nu-1}$ avec $\langle y_\nu, x_\nu \rangle = 0$.

Définissant les $(N \times \nu)$ matrices

$$X = [x_0 \ x_1 \ \dots \ x_{\nu-1}], \quad Y = [y_0 \ y_1 \ \dots \ y_{\nu-1}]$$

et les matrices de Hessenberg inférieures d'ordre ν

$$T = [\tau_{j,k}], \quad S = [\sigma_{j,k}],$$

nous pouvons écrire les expressions (1.5) et (1.6) sous forme matricielle

$$\begin{aligned} AX &= XT + \tau_{\nu,\nu-1} x_\nu e_\nu^T, \\ \tilde{A}Y &= YS + \sigma_{\nu,\nu-1} y_\nu e_\nu^T \end{aligned} \quad (1.7)$$

où e_ν est le ν -ième vecteur de la base canonique. Les relations de biorthogonalité (1.2) deviennent

$$\tilde{Y}X = D = \text{diag}(\delta_0, \delta_1, \dots, \delta_{\nu-1})$$

de sorte que, par (1.7),

$$\begin{aligned} \tilde{Y}AX &= DT, \\ \tilde{X}\tilde{A}Y &= \tilde{D}S, \end{aligned}$$

soit

$$DT = \tilde{S}D. \quad (1.8)$$

Dans le membre de gauche, la matrice est de type Hessenberg inférieur tandis que, dans celui de droite, elle est de type Hessenberg supérieur : les deux matrices T et S sont donc tridiagonales

$$T = \begin{pmatrix} \alpha_0 & \beta_1 & & & \\ \gamma_0 & \alpha_1 & \beta_2 & & \\ & \gamma_1 & \alpha_2 & \ddots & \\ & & \ddots & \ddots & \beta_{\nu-1} \\ & & & \gamma_{\nu-2} & \alpha_{\nu-1} \end{pmatrix}, \quad S = \begin{pmatrix} a_0 & b_1 & & & \\ c_0 & a_1 & b_2 & & \\ & c_1 & a_2 & \ddots & \\ & & \ddots & \ddots & b_{\nu-1} \\ & & & c_{\nu-2} & a_{\nu-1} \end{pmatrix}.$$

Par (1.8), nous avons

$$a_n = \overline{\alpha_n}$$

et

$$\delta_{n-1}\beta_n = \delta_n \overline{c_{n-1}}, \quad \delta_{n-1}\overline{b_n} = \delta_n \gamma_{n-1}, \quad (1.9)$$

soit

$$\beta_n \gamma_{n-1} = \overline{b_n c_{n-1}}.$$

Nous pouvons satisfaire à cette dernière relation en posant

$$b_n = \overline{\beta_n}, \quad c_{n-1} = \overline{\gamma_{n-1}}$$

soit

$$S = \overline{T}.$$

Par ce choix de S , (1.9) donne la relation

$$\delta_n = \frac{\delta_{n-1}\beta_n}{\gamma_{n-1}}. \quad (1.10)$$

En résumé, les relations (1.7) s'écrivent, pour $n = 0, 1, \dots, \nu-1$,

$$Ax_n = \beta_n x_{n-1} + \alpha_n x_n + \gamma_n x_{n+1}, \quad (1.11)$$

$$\tilde{A}y_n = \overline{\beta_n} y_{n-1} + \overline{\alpha_n} y_n + \overline{\gamma_n} y_{n+1} \quad (1.12)$$

avec $\beta_0 = 0$, $x_{-1} = y_{-1} = 0$.

Effectuant le produit scalaire de Ax_n avec y_{n-1} et y_n , on obtient :

$$\begin{aligned} \langle y_{n-1}, Ax_n \rangle &= \beta_n \delta_{n-1}, \\ \langle y_n, Ax_n \rangle &= \alpha_n \delta_n. \end{aligned} \quad (1.13)$$

Utilisant (1.10) à (1.13), nous obtenons la forme suivante de l'algorithme de Lanczos.

1.2 L'algorithme de biorthogonalisation (BO).

Choisir $x_0, y_0 \in \mathbb{C}^N$ tels que $\delta_0 = \langle y_0, x_0 \rangle \neq 0$ et poser $\beta_0 := 0$
Calculer pour $n = 0, 1, \dots$

$$\begin{aligned} \alpha_n &= \frac{\langle y_n, Ax_n \rangle}{\delta_n}, \\ \beta_n &= \frac{\langle y_{n-1}, Ax_n \rangle}{\delta_{n-1}} = \frac{\gamma_{n-1} \delta_n}{\delta_{n-1}} \quad (n \geq 0), \end{aligned} \quad (1.14)$$

$$x_{n+1}^* = Ax_n - x_n \alpha_n - x_{n-1} \beta_n, \quad (1.15)$$

$$y_{n+1}^* = \tilde{A}y_n - y_n \overline{\alpha_n} - y_{n-1} \overline{\beta_n}, \quad (1.16)$$

$$\delta_{n+1}^* = \langle y_{n+1}^*, x_{n+1}^* \rangle. \quad (1.17)$$

Si $\delta_{n+1}^* = 0$, choisir $\gamma_n \neq 0$ quelconque et poser

$$\nu := n + 1, \quad x_\nu := \frac{x_\nu^*}{\gamma_n}, \quad y_\nu := \frac{y_\nu^*}{\gamma_n}, \quad \delta_{n+1} := 0 \quad (1.18)$$

et arrêter. Sinon, choisir γ_n et δ_{n+1} tels que

$$\gamma_n^2 \delta_{n+1} = \delta_{n+1}^*. \quad (1.19)$$

Calculer

$$\begin{aligned} x_{n+1} &= \frac{x_{n+1}^*}{\gamma_n}, \\ y_{n+1} &= \frac{y_{n+1}^*}{\gamma_n}. \end{aligned} \tag{1.20}$$

Si $x_{n+1} \neq 0$ et $y_{n+1} \neq 0$, procéder à l'étape suivante.

Théorème 1.2 1. Si, $\forall m, n = 0, 1, \dots, \nu - 1$, les conditions (1.1) et (1.2) sont satisfaites et si $S = \overline{T}$, alors, les formules de récurrence de l'algorithme BO sont vérifiées pour $n = 0, 1, \dots, \nu - 2$. Pour $n = \nu - 1$, un breakdown peut se produire par suite de $\delta_\nu^* = 0$.

Les conditions (1.1) et (1.2) déterminent les suites $\{x_n\}_{n=0}^{\nu-1}$ et $\{y_n\}_{n=0}^{\nu-1}$ univoquement, à un facteur scalaire près.

2. Inversément, l'algorithme BO génère les séquences $\{x_n\}_{n=0}^{\nu-1}$ et $\{y_n\}_{n=0}^{\nu-1}$ satisfaisant à (1.1) et (1.2).

Preuve. La première partie du théorème est immédiate puisqu'on a construit l'algorithme à partir des conditions (1.1) et (1.2).

Pour la seconde partie, nous disposons des séquences $\{x_n\}_{n=0}^{\nu-1}$ et $\{y_n\}_{n=0}^{\nu-1}$ générées par l'algorithme BO. Procédons par récurrence. On suppose que (1.1) et (1.2) soient vérifiées pour $n = 0, \dots, k$ ($\leq \nu$) (pour $k = 0$, c'est immédiat). Regardons ce qui se passe lorsque $n = k + 1$. La condition (1.1) est satisfaite par (1.15) et (1.16);

Pour la condition (1.2), grâce à (1.15), (1.17) et (1.20), nous avons :

$$\begin{aligned} \langle y_m, x_{k+1} \rangle &= \tilde{y}_m \frac{(Ax_k - x_k \alpha_k - x_{k-1} \beta_k)}{\gamma_k} \\ &= \frac{\gamma_m \tilde{y}_{m+1} x_k + \alpha_k \tilde{y}_m x_k + \beta_m \tilde{y}_{m-1} x_k - \tilde{y}_m x_k \alpha_k - \tilde{y}_m x_{k-1} \beta_k}{\gamma_k}. \end{aligned}$$

Si $m \leq k - 2$, tous les termes du membre de droite s'annulent.

Si $m = k - 1$, $\langle y_{k-1}, x_{k+1} \rangle = (\beta_k \delta_{k-1} - 0 - \beta_k \delta_{k-1}) = 0$.

Si $m = k$, $\langle y_k, x_{k+1} \rangle = \frac{\alpha_k \delta_k - \alpha_k \delta_k}{\gamma_k} = 0$.

Enfin, si $m = k + 1$, $\langle y_{k+1}, x_{k+1} \rangle = \frac{1}{\gamma_k^2} \delta_{k+1}$. □

1.3 Les différents "breakdowns".

Il existe deux types de "breakdowns".

- $x_\nu = 0$ ou $y_\nu = 0$,

- $x_\nu \neq 0$ et $y_\nu \neq 0$, mais $\langle y_\nu, x_\nu \rangle = 0$.

Dans le premier cas, nous parlerons d'arrêt prématuré, dans le second, de "breakdown sérieux" [5].

1.3.1 Arrêt prématuré.

Introduisons avant toute chose la notion de polynôme minimal d'une matrice A [6] et [9].

Définition 1.3 Soit la séquence de matrices $\{I, A, A^2, \dots, A^s\}$ telle que :

$$A^s + m_{s-1}A^{s-1} + \dots + m_0I = 0$$

où s est le degré minimal pour lequel la séquence de matrices $\{I, A, A^2, \dots, A^s\}$ est linéairement indépendante.

Le polynôme $m(\lambda)$ correspondant au membre de gauche de l'équation ci-dessus est appelé polynôme minimal de la matrice A .

Proposition 1.4 Si $x_\nu = 0$ ($y_\nu = 0$), le sous-espace $K_{\nu-1}$ ($L_{\nu-1}$) est invariant de dimension ν de A (\tilde{A}).

Preuve. Par (1.7), $x_\nu = 0$ entraîne $AX = XT$ où $X = [x_0 \dots x_{\nu-1}]$ et $\text{rang}(X) = \nu$.

Or, $K_{\nu-1} = \text{Im}_A(X) = [Ax_0 \dots Ax_{\nu-1}]$. Pour $x \in K_{\nu-1}$, on a $x = Xy$ et donc, $AX = AXy = XTy$, soit $Xz \in K_{\nu-1}$ où $z = Ty$. \square

La proposition suivante nous indique qu'en fait, le "breakdown" prématuré se produira au plus tard quand ν vaudra le degré du polynôme minimal de A .

Proposition 1.5 La dimension du sous-espace de Krylov de A (\tilde{A}) est au plus égale au degré du polynôme minimal de A (\tilde{A}).

Preuve. Soit $p_m(A)$, le polynôme minimal de A , i.e. $p_m(A) = \sum_{j=0}^m \alpha_j A^j$ ($\alpha_m = 1$) tel que $p_m(A) = 0$. On a $A^m = -\sum_{j=0}^{m-1} \alpha_j A^j$ ou encore, $A^m x_0 = -\sum_{j=0}^{m-1} \alpha_j A^j x_0$ de sorte que $K_m = K_{m-1}$. \square

Nous avons de plus la propriété suivante :

Propriété 1.6 *Le spectre de la matrice T est un sous-ensemble du spectre de la matrice A .*

Preuve. Soit λ , une valeur propre appartenant au spectre de la matrice T où

$$sp(T) \equiv \{\lambda \in \mathbb{C} \mid T - \lambda I \text{ non inversible}\}.$$

Nous avons donc

$$Tu = \lambda u$$

avec $u \neq 0$. Par conséquent,

$$A(Xu) = XT u = \lambda(Xu)$$

où $(A - \lambda I)Xu$ est non inversible ($u, X \neq 0$). λ est donc une valeur propre appartenant au spectre de la matrice A . Par conséquent, $sp(T) \subseteq sp(A)$. \square

Si le "breakdown" prématuré se manifeste avant que ν n'atteigne le degré du polynôme minimal de A , il y a lieu de redémarrer l'algorithme avec une paire (x_ν, y_ν) biorthogonale aux vecteurs construits jusque là.

Si $x_\nu = y_\nu = 0$, une possibilité pour satisfaire (1.14) est de prendre $\gamma_{\nu-1} = \beta_\nu$ de telle manière que la relation (1.7) soit vérifiée même au-delà de l'indice ν . Dans ce cas, après plusieurs redémarrages possibles, l'algorithme doit se terminer avec $x_N = y_N = 0$ et $\nu = N$. Il suit donc que

$$\begin{aligned} AX &= XT, \\ \tilde{A}Y &= YS \quad (= Y\bar{T}), \\ \tilde{Y}X &= D_\delta = \text{diag}(\delta_0, \delta_1, \delta_2, \dots, \delta_{N-1}), \end{aligned} \tag{1.21}$$

où toutes les matrices sont carrées d'ordre N .

Certains éléments ($\gamma_j = \beta_{j+1}$) de la matrice tridiagonale T peuvent être nuls suite aux redémarrages. X et Y étant singulières, on a que T est similaire à A et S à \tilde{A} .

Ceci amène donc la proposition suivante.

Proposition 1.7 *Pour toute matrice A , il existe une matrice tridiagonale T avec $\gamma_j = 0$ si et seulement si $\beta_{j+1} = 0$*

Rappelons-nous que, pour satisfaire (1.9), on avait choisi $S = \overline{T}$, une autre possibilité [1] aurait pu être: $S = \tilde{T}$.

On aurait obtenu à la place de (1.21):

$$\begin{aligned} AX &= XT, \\ \tilde{A}Y &= Y\tilde{T}. \end{aligned}$$

Tout comme le choix $S = \overline{T}$ déterminait univoquement les séquences $\{x_n\}$ et $\{y_n\}$, le choix $S = \tilde{T}$ déterminera la même séquence $\{x_n\}$ et une séquence $\{y_n\}$ identique également à la précédente, aux coefficients près.

1.3.2 "Breakdown" sérieux.

Si la mise en marche de l'algorithme BO s'effectue avec une paire (x_0, y_0) telle que $x_0 \neq y_0$, l'algorithme peut présenter un second type de "breakdown". En effet, pour certains n , il se peut que tout x_n et y_n satisfaisant à (1.3) et (1.4) soient orthogonaux l'un à l'autre de telle manière qu'on ne puisse satisfaire à (1.2).

1.3.3 Remarque.

Des problèmes de précision peuvent apparaître. En effet, l'emploi du processus BO est, en pratique rendu hasardeux à cause des problèmes d'arrondis. A la n -ième étape de l'algorithme, les constantes $\alpha_n, \beta_n, \gamma_n, \delta_n$ et les vecteurs x_{n+1} et y_{n+1} sont déterminés de telle manière que :

$$\begin{aligned} x_{n+1} &\perp y_n, \\ x_{n+1} &\perp y_{n-1}, \\ \langle y_{n+1}, x_{n+1} \rangle &= \delta_{n+1}. \end{aligned}$$

Quand n devient grand, la précision de la biorthogonalité se perd (rappelons que la projection orthogonale réduit nécessairement la taille d'un vecteur et donc, par conséquent, sa précision relative).

1.4 Algorithme de résolution de systèmes d'équations linéaires.

Soit, à résoudre :

$$Az = b \tag{1.22}$$

où A est non singulière.

Notons :

- z^* : la solution de $Az = b$,
- z_0 : l'approximation initiale de la solution,
- $x_0 = b - Az_0$: le résidu correspondant,
- K_n : le n -ième espace de Krylov généré par x_0 .

Dans la méthode du gradient conjugué, la n -ième itération consiste à minimiser la fonction quadratique :

$$z \mapsto \langle z^* - z, A(z^* - z) \rangle$$

parmi tous les $z \in z_0 + K_{n-1}$ [6]. La solution optimale aura comme propriété que les résidus $x_n = b - Az_n$ forment une séquence orthogonale. Cela entraîne $x_\nu = 0$ pour $\nu \leq N$ et par conséquent, $z_\nu = z^*$.

Au cours de cet algorithme, nous construirons donc une séquence biorthogonale :

$$x_n \perp L_{n-1}, \quad x_n = b - Az_n.$$

Proposition 1.8 *La matrice X vérifie la relation $AX = XH$ si et seulement si la somme des éléments dans chaque colonne est nulle.*

Preuve. Par analogie avec (1.7), les n premiers pas de la récurrence pour les résidus x_k peuvent toujours être écrits comme $AX_n = X_{n+1}H_{n+1,n}$ où $H_{n+1,n}$ est une $(n+1) \times n$ matrice de Hessenberg.

Un simple calcul donne :

$$X_n = A^{-1}b[1 \dots 1]H_{n+1,n} - Z_{n+1}H_{n+1,n}$$

où $Z_{n+1} = [z_0, z_1, \dots, z_n] \in \mathbb{C}^{N \times (n+1)}$. Nous obtenons donc que $A^{-1}b[1 \dots 1]H_{n+1,n} = 0$ si et seulement si la somme de chaque colonne de $H_{n+1,n}$ est égale à 0. \square

Dans notre cas, $H = T$; prendre $\gamma_n = -\alpha_n - \beta_n$.

Pour résoudre $Az = b$, poser z_0 une approximation initiale telle que $x_0 = b - Az_0$. Choisir y_0 tel que $\langle y_0, x_0 \rangle \neq 0$. Ensuite, appliquons l'algorithme de la page 10 avec

$$\gamma_n = -\alpha_n - \beta_n \tag{1.23}$$

de telle manière que (1.17) et (1.20) se simplifient :

$$\begin{aligned}x_{n+1} &= \frac{Ax_n - x_n\alpha_n - x_{n-1}\beta_n}{\gamma_n}, \\y_{n+1} &= \frac{\tilde{A}y_n - y_n\overline{\alpha_n} - y_{n-1}\overline{\beta_n}}{\overline{\gamma_n}}, \\ \delta_{n+1} &= \langle y_{n+1}, x_{n+1} \rangle.\end{aligned}$$

Calculons en plus les vecteurs :

$$z_{n+1} = \frac{-(x_n + z_n\alpha_n + z_{n-1}\beta_n)}{\gamma_n}. \quad (1.24)$$

Si le résidu est nul ($x_{n+1} = 0$), le processus est terminé et z_{n+1} est la solution. Si $x_{n+1} \neq 0$ mais $\delta_{n+1} = 0$, un "breakdown" sérieux surviendra (dans ce cas, nous poserons $\nu = n + 1$). Cet algorithme peut également présenter un nouveau type de "breakdown" : le "breakdown" de normalisation. Celui-ci provient du fait que γ_n peut être nul. Dans ce cas, nous poserons $\nu = n$.

Lemme 1.9 *Dans cet algorithme, le vecteur x_n est le n -ième résidu $b - Az_n = x_n$ ($n = 0, 1, \dots, \nu$)*

Preuve. Démontrons ce lemme par récurrence. Nous supposons que $b - Az_k = x_k$ ($k = 1, \dots, n$) ($k = 0$ étant un cas immédiat). Pour $k = n + 1$, en vertu de (1.17), (1.20), (1.23), nous avons :

$$b - Az_{n+1} = \frac{b + (Ax_n + Az_n\alpha_n + Az_{n-1}\beta_n)}{\gamma_n} = x_{n+1}. \quad \square$$

Grâce à $z = [z_0, \dots, z_{\nu-1}]$, la récurrence (1.24) pour z_n peut être écrite comme :

$$X = -ZT - z_\nu \gamma_{\nu-1} e_\nu^T$$

par un raisonnement identique à celui conduisant à (1.7).

La condition somme (1.23) de chaque colonne peut être réécrite de cette manière

$$[1, 1, \dots, 1]T = \gamma_{\nu-1} e_\nu^T.$$

Chapitre 2

Les polynômes orthogonaux formels et les approximants de Padé.

Le but de ce chapitre est d'énoncer quelques théorèmes et propriétés permettant d'établir un lien entre les polynômes orthogonaux formels et les approximants de Padé [2]. Nous procéderons comme suit; après la définition des deux concepts, nous énoncerons quelques propriétés pour, en fin de section, conclure sur un théorème reprenant les résultats principaux.

Soit donnée une séquence complexe $\{\psi_k\}_{k=0}^{\infty}$.
Posons la fonctionnelle linéaire $\phi : \mathcal{P} \rightarrow \mathbb{C}$ définie par les valeurs que prend la base monomiale :

$$\phi(z^k) := \psi_k \quad (k \in \mathbb{N}). \quad (2.1)$$

Définition 2.1 $P_n \in \mathcal{P}_n$ ($n \geq 1$) est un n -ième polynôme orthogonal formel normalisé vrai par rapport à ϕ si P_n est monique ($P_0(z) := 1$) de degré n et si on a :

$$\phi(pP_n) = 0, \quad \forall p \in \mathcal{P}_{n-1}. \quad (2.2)$$

De plus, si P_n est unique, nous dirons qu'il est régulier; sinon, nous dirons qu'il est singulier.

Considérons la série formelle :

$$f(z) := \sum_{k=0}^{\infty} \psi_k z^k \quad (2.3)$$

Propriété 2.2 Etant donné que $z^n P_n(z^{-1})$ est un polynôme, le produit :

$$f(z)z^n P_n(z^{-1})$$

est une série formelle. Soit,

$$P_n(z) := \sum_{j=0}^n \pi_{j,n} z^j, \quad (2.4)$$

Alors,

$$f(z)z^n P_n(z^{-1}) = \sum_{i=0}^{\infty} \left(\sum_{j=0}^n \psi_{i+j-n} \pi_{j,n} \right) z^i. \quad (2.5)$$

Le polynôme orthogonal P_n est caractérisé par le système homogène de n équations linéaires à $(n+1)$ inconnues suivant :

$$\sum_{j=0}^n \psi_{i+j-n} \pi_{j,n} = 0 \quad (i = n, \dots, 2n-1). \quad (2.6)$$

En effet, par (2.2), nous savons que $\phi(z^k P_n) = 0$ pour $k = 0, \dots, n-1$ où $P_n(z) = \sum_{j=0}^n \pi_{j,n} z^j$. La linéarité de ϕ nous donne

$$\sum_{j=0}^n \pi_{j,n} \phi(z^{k+j}) = 0.$$

Enfin, par la définition (2.1), nous obtenons la relation (2.6). Si nous réécrivons le système complet, nous avons :

$$\begin{pmatrix} \psi_0 & \psi_1 & \dots & \psi_{n-1} & \psi_n \\ \psi_1 & \psi_2 & \dots & \psi_n & \psi_{n+1} \\ \dots & \dots & \dots & \dots & \dots \\ \psi_{n-1} & \psi_n & \dots & \psi_{2n-2} & \psi_{2n-1} \end{pmatrix} \begin{pmatrix} \pi_{0,n} \\ \pi_{1,n} \\ \dots \\ \pi_{n-1,n} \\ \pi_{n,n} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \\ 0 \end{pmatrix}.$$

Nous voyons donc que les coefficients de z^i dans (2.5) sont nuls pour $i = n, \dots, 2n-1$. (2.5) devient donc :

$$f(z)z^n P_n(z^{-1}) = p_{n-1,n}(z) + \mathcal{O}_+(z^{2n}) \quad (2.7)$$

où

$$p_{n-1,n}(z) = \sum_{i=0}^{n-1} \left(\sum_{j=0}^n \psi_{i+j-n} \pi_{j,n} \right) z^i, \quad (2.8)$$

r_{00}	r_{01}	r_{02}	r_{03}	-----
r_{10}	r_{11}	r_{12}	r_{13}	-----
r_{20}	r_{21}	r_{22}	r_{23}	-----
r_{30}	r_{31}	r_{32}	r_{33}	-----
\vdots	\vdots	\vdots	\vdots	

FIG. 2.1 - Table de Padé

est un polynôme de degré au plus $(n-1)$ et où $\mathcal{O}_+(z^k)$ regroupe tous les termes de degré plus grand ou égal à k .

Définition 2.3 Pour $f \in \mathcal{P}$, $m \in \mathbb{N}$ et $n \in \mathbb{N}$, une forme de Padé (m, n) de f est une paire $(p, q) \in \mathcal{P}_m \times \mathcal{P}_n$ avec $(p, q) \neq (0, 0)$ pour laquelle

$$f(z)q(z) - p(z) = \mathcal{O}_+(z^{m+n+1}). \quad (2.9)$$

La fonction rationnelle correspondante $r_{m,n} = \frac{p}{q}$ est appelée approximant (m, n) de Padé. Si r est aussi solution de

$$f(z) - r(z) = \mathcal{O}_+(z^{m+n+1}).$$

Nous l'appelons interpolant vrai (m, n) de Padé. Dans le cas contraire, r est un approximant déficient de Padé.

Définition 2.4 La table de Padé d'une série formelle

$$C(z) = \sum_{m=0}^{\infty} c_m z^m$$

est un tableau doublement infini de fonctions rationnelles

$$r_{m,n}(z) = \frac{p_{m,n}(z)}{q_{m,n}(z)} = \frac{a_0 + a_1 z + \dots + a_m z^m}{b_0 + b_1 z + \dots + b_n z^n}$$

déterminées de telle manière que le développement de Mac Laurin de $r_{m,n}$ s'identifie le plus loin possible avec $C(z)$ [2].

Graphiquement, on représente le table comme à la figure (2.1).

Propriété 2.5 (i) En comparant (2.7) et (2.9) et en posant

$$q_{n-1,n}(z) := z^n P_n(z^{-1}), \quad (2.10)$$

la paire $(p_{n-1,n}, q_{n-1,n}) \in \mathcal{P}_{n-1} \times \mathcal{P}_n$ est une forme de Padé $(n-1, n)$ de f .

(ii) Si (p, q) n'est pas une forme de Padé $(n-1, n)$ de f , alors, en posant $P_n(z) := z^n q(z^{-1})$, (2.9) implique l'expression (2.7) qui, lorsqu'on se restreint à $P_n \in \mathcal{P}_n$, est équivalent à (2.2).

Le point (ii) de cette propriété entraîne la définition suivante.

Définition 2.6 P_n est un vrai polynôme orthogonal formel normalisé par rapport à ϕ si et seulement si il a un degré exact n et est monique.

Etant donnés f , m et n , chaque forme (m, n) de Padé peut être construite en résolvant le système linéaire (2.6) avec $n-1 = m$, i.e., $m-n+1 = 0$, et en définissant, pour une solution non triviale, $\{\pi_{i,j,n}\}_{j=0}^n$, de (2.6)

$$\begin{aligned} q(z) &:= \sum_{j=0}^n \pi_{j,n} z^{n-j}, \\ p(z) &:= \sum_{i=0}^m \left(\sum_{j=0}^n \psi_{i+j-n} \pi_{j,n} \right) z^i. \end{aligned} \quad (2.11)$$

Pour la suite, le résultat de base le plus important est celui concernant l'allure générale des formes de Padé [2]:

Théorème 2.7 Pour $f \in \mathcal{P}$, $m \in \mathbb{N}$, $n \in \mathbb{N}$, la solution générale $(p, q) \in \mathcal{P}_m \times \mathcal{P}_n$ de (2.9) est

$$(p, q) = (z^\sigma \hat{p}_{m,n} \omega, z^\sigma \hat{q}_{m,n} \omega) \quad (2.12)$$

où $\hat{p}_{m,n} \in \mathcal{P}_m$ et $\hat{q}_{m,n} \in \mathcal{P}_n$ sont déterminés univoquement. On a $\hat{q}_{m,n}(0) = 1$, $\sigma := \sigma_{m,n}$ est un entier fixé où

$$0 \leq \sigma \leq \delta := \min\{m - \partial \hat{p}_{m,n}; n - \partial \hat{q}_{m,n}\},$$

δ étant appelé la déficiencia de $\frac{p}{q}$. Enfin, $\omega \in \mathcal{P}_{\delta-\sigma}$ est arbitraire.

Pour démontrer ce théorème, nous devons établir un résultat préliminaire [3]:

Propriété 2.8 Soit $f \in \mathcal{P}$ et $(m, n) \in \mathbb{N} \times \mathbb{N}$. Il existe des paires $(p, q) \in \mathcal{P}_m \times \mathcal{P}_n$ vérifiant (2.9) et elles sont uniques dans le sens où $\frac{p}{q}$ définit une fraction unique $r \in \mathcal{R}_{m,n}$.

Preuve. L'existence d'une telle paire (p, q) est immédiate car (2.9) est équivalent à un système dans lequel $(m + n + 1)$ fonctionnelles linéaires agissent sur (p, q) . Cela correspond donc à un système linéaire homogène de $(m + n + 1)$ équations à $(m + n + 2)$ inconnues. Il s'agit par conséquent d'un système linéaire homogène sous-déterminé admettant toujours une solution non nulle.

Pour l'unicité, supposons par l'absurde qu'on ait deux solutions (p, q) et (\tilde{p}, \tilde{q}) . Par (2.9), $f q - p = \mathcal{O}_+(z^{m+n+1})$ et $f \tilde{q} - \tilde{p} = \mathcal{O}_+(z^{m+n+1})$. Cela entraîne $f(\tilde{p} q - \tilde{q} p) = \mathcal{O}_+(z^{m+n+1})$. f étant quelconque dans \mathcal{P} , cela donne $\tilde{p} q - \tilde{q} p = \mathcal{O}_+(z^{m+n+1})$. Or, $\tilde{p} q - \tilde{q} p \in \mathcal{P}_{m+n}$; il y a donc une contradiction. \square

Cette propriété établie, nous pouvons démontrer le théorème 2.7.

Preuve (théorème 2.7). Soit $(p, q) \in \mathcal{P}_m \times \mathcal{P}_n$ une solution quelconque de (2.9). Soit δ , la déficience de $\frac{p}{q}$.

- On peut écrire $(p, q) = (z^\sigma \hat{p} \omega, z^\sigma \hat{q} \omega)$ où

* z^σ est déterminé en choisissant σ pour mettre en évidence un exposant commun à p et à q . σ sera choisi de sorte que ce soit le plus petit exposant pour lequel $(z^\sigma \hat{p}, z^\sigma \hat{q})$ est encore une solution.

* ω est un polynôme dans lequel on regroupe tous les termes communs de p et q restant après mise en évidence de z^σ .

On a $\partial \omega + \sigma \leq \delta$ et donc, $\partial \omega \leq \delta - \sigma$, i.e. $\omega \in \mathcal{P}_{\delta-\sigma}$.

- \hat{p} et \hat{q} sont uniques au sens de la propriété 2.8. z^σ est unique aussi car σ peut être caractérisé comme le degré minimal pour lequel $f \hat{q} z^\sigma - \hat{p} z^\sigma = \mathcal{O}_+(z^{m+n+1})$, i.e.,

$$f \hat{q} - \hat{p} = \mathcal{O}_+(z^{m+n+1-\sigma}). \quad (2.13)$$

- En multipliant (2.13) par $z^\sigma \omega$ (où $\omega \in \mathcal{P}_{\delta-\sigma}$ est arbitraire), on a une solution de (2.9). \square

Plusieurs corollaires peuvent être déduits du théorème 2.7.

Corollaire 2.9 *L'approximant (m, n) de Padé $r_{m,n} = \frac{\hat{p}_{m,n}}{\hat{q}_{m,n}}$ est un vrai interpolant de Padé si et seulement si $\sigma = 0$ dans (2.12).*

Corollaire 2.10 *Le système linéaire homogène (2.6) a un rang $n - \delta + \sigma$, i.e. un rang de déficience $(\delta - \sigma)$.*

Corollaire 2.11 (théorème de caractérisation) :

Une fonction rationnelle $r = \frac{p}{q} \in \mathcal{R}_{m,n}$ avec $\delta = \min\{m - \partial p, n - \partial q\}$ est l'approximant (m, n) de Padé de f si et seulement si

$$f(z) - r(z) = \mathcal{O}_+(z^{m+n+1-\delta}).$$

Théorème 2.12 (théorème de structure en bloc) [2]

(i) Une fonction rationnelle $r \in \mathcal{R}_{\mu,\nu}$ ($r \neq 0$) de type exact (μ, ν) est un approximant de Padé de $f \notin \mathcal{R}_{\mu,\nu}$ si et seulement si

$$f(z) - r(z) \equiv \mathcal{O}_+(z^{\mu+\nu+\Delta+1}) \quad (2.14)$$

avec $\Delta \in \mathbb{N}$.

Si cette dernière égalité est vérifiée, $r_{m,n}$ est l'approximant de Padé (m, n) de f pour toute paire (m, n) dans le carré

$$\{(m, n); \mu \leq m \leq \mu + \Delta, \nu \leq n \leq \nu + \Delta\}$$

et pour aucune autre paire.

Dans (2.12), on voit alors que $\partial \hat{p}_{m,n} = \mu$, $\partial \hat{q}_{m,n} = \nu$, $\delta = \min\{m - \mu, n - \nu\}$ et

$$\sigma = \max\{0, m + n - \mu - \nu - \Delta\} \quad (2.15)$$

de telle manière que

$$\delta - \sigma = \min\{m - \mu, n - \nu, \mu + \Delta - m, \nu + \Delta - n\}.$$

(ii) Si f est elle-même rationnelle, de type exact (μ, ν) , alors $r_{m,n} = f$ pour tout (m, n) dans le quadrant $\{(m, n); m \geq \mu, n \geq \nu\}$. De plus, dans (2.12), $\partial \hat{p}_{m,n} = \mu$, $\partial \hat{q}_{m,n} = \nu$, $\delta = \min\{m - \mu, n - \nu\}$ et $\sigma = 0$.

(iii) La fonction $r = 0$ de type exact $(-\infty, 0)$ est un approximant de Padé de $f \neq 0$ si et seulement si $f(z) = \mathcal{O}_+(z^k)$ pour certains $k > 0$; alors $r_{m,n}$ est identiquement nul pour tout (m, n) dans le demi-quadrant $\{(m, n); 0 \leq m < k, n \in \mathbb{N}\}$ (voir figure (2.2)).

Ceci entraîne qu'en (2.12), $\hat{p}_{m,n} = 0$, $\hat{q}_{m,n} = 1$, $\delta = n$ et $\sigma = \max\{0, m + n + 1 - k\}$.

Corollaire 2.13 (définition par optimalité) :

L'approximant de Padé (m, n) $r_{m,n}$ est, parmi tous les éléments r de $\mathcal{R}_{m,n}$, le seul pour lequel l'égalité

$$f(z) - r(z) = \mathcal{O}_+(z^k)$$

est vérifiée pour la plus grande valeur de k possible ($k \in \mathbb{Z}$).

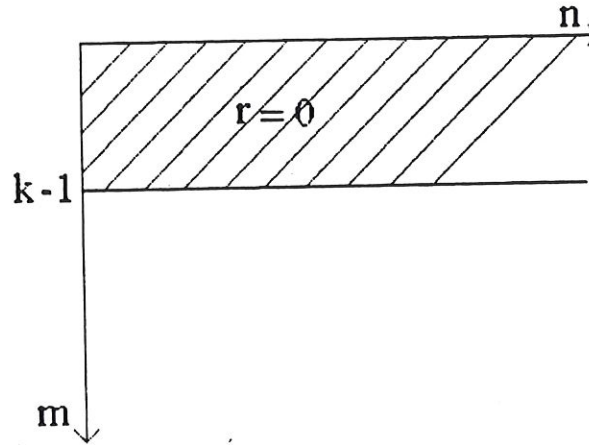


FIG. 2.2 - Illustration du point (iii) du théorème de structure en bloc

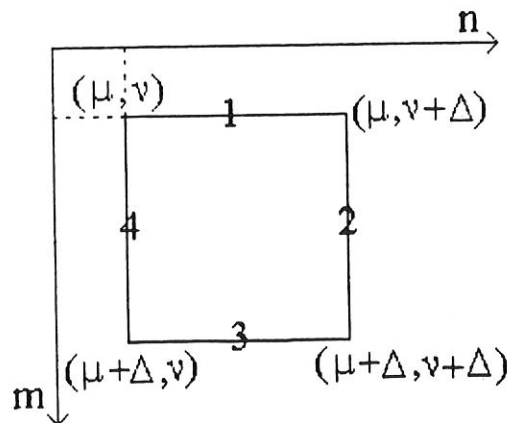


FIG. 2.3 - Sur le côté 1, $\delta = \sigma = 0$ par définition de δ et de $\delta - \sigma$ où $m = \mu$ - Sur le côté 2, $\delta = \sigma$ par définition de $\delta - \sigma$ où $n = \nu + \Delta$ - Sur le côté 3, $\delta = \sigma$ par définition de $\delta - \sigma$ où $m = \mu + \Delta$ - Sur le côté 4, $\delta = \sigma = 0$ par définition de δ et de $\delta - \sigma$ où $n = \nu$.

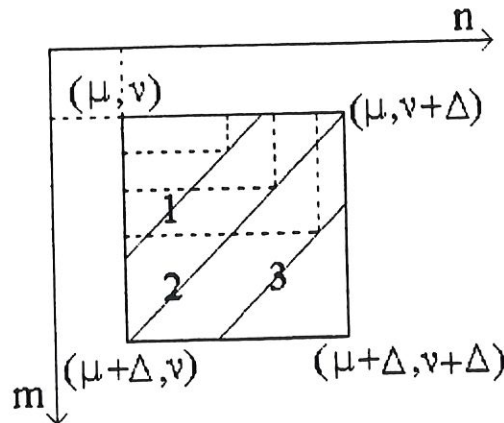


FIG. 2.4 - Sur l'antidiagonale 1, $m+n = \mu + \nu + \Delta$ et donc, $\sigma = 0$ - Sur l'antidiagonale 2, $m+n \leq \mu + \nu + \Delta$ et donc, $\sigma = 0$ - Sur l'antidiagonale 3, $m+n \geq \mu + \nu + \Delta$ et donc, $\sigma \geq 0$.

Selon le théorème de structure en bloc, la table de Padé, définie en (2.4) qui contient l'approximant de Padé (m, n) de f à l'entrée (m, n) , est composée de blocs carrés distincts de fonctions rationnelles (2.14) [3].

Pour chaque bloc, la position (μ, ν) du coin supérieur gauche indique le type exact de l'approximant. Sur le bord du bloc (voir figure (2.3)), on a $\delta = \sigma$, ce qui signifie que la forme de Padé est unique à un facteur scalaire commun près. Sur et au-dessus de l'antidiagonale (voir figure (2.4)), $\sigma = 0$ de telle manière que la fonction r soit un vrai interpolant de Padé à cet endroit puisque, en-dessous de l'antidiagonale, r est un approximant déficient. Il y a deux exceptions à ces blocs carrés finis : si la fonction est nulle, alors son bloc est une demi-bande. Si f est elle-même rationnelle, il y a un "bloc carré infini", puisque dans le quadrant tout entier, $\{(m, n); m \geq \mu, n \geq \nu\}$, l'approximant de Padé est $r = f$.

Revenons aux polynômes orthogonaux formels définis en (2.1). Tout comme nous l'avons vu au début de ce chapitre, on peut établir une relation entre un n -ième polynôme orthogonal formel satisfaisant à (2.2) et l'approximant de Padé $(n-1, n)$ de f . Puisque $P_n (n \geq 1)$ est supposé avoir un degré exact n , nous avons besoin, en vue de satisfaire à (2.10) d'une forme de Padé $(n-1, n), (p, q)$ de f avec $q(0) \neq 0$.

P_n existe si et seulement si $\sigma = 0$ dans le théorème (2.7) et il y a un polynôme P_n unique monique si et seulement si $\sigma = \delta = 0$. Notons que δ est la quantité de laquelle n surpasse le plus petit indice suivant pour lequel P_n est unique, i.e. régulier.

En résumé, nous avons :

Théorème 2.14 *Pour $f \in \mathcal{P}, n \in \mathbb{N}^+$, il existe un vrai polynôme orthogonal formel P_n si et seulement si la paire d'indice $(m, n) = (n-1, n)$ se situe soit dans un bloc infini, non nul, soit sur ou au-dessus de l'antidiagonale d'un bloc carré fini ou bloc nul de la table de Padé de f , i.e. si et seulement si f a un interpolant vrai (m, n) de Padé. P_n a alors la forme générale :*

$$P_n(z) = \hat{P}_n(z)W(z) = z^n \hat{q}_{m,n}(z^{-1})\omega(z^{-1}) \quad (2.16)$$

où $\hat{P}_n(z) = z^{n-\delta} \hat{q}_{m,n}(z^{-1})$ est monique de degré $(n - \delta)$ et $\hat{q}_{m,n}$ est le dénominateur normalisé de l'approximant (m, n) de Padé.

$W(z) = z^\delta \omega(z^{-1})$ est un polynôme arbitraire monique de degré exact δ . De plus, $\hat{P}_n = P_{n_j}$, où on note n_j le plus grand entier plus petit ou égal à n pour lequel un polynôme formel orthogonal régulier par rapport à ϕ existe (donc, $\delta = n - n_j$).

P_n est régulier (i.e. unique) si et seulement si la paire (m, n) est située sur le bord supérieur ou à gauche d'un bloc carré fini ou infini de la table de Padé. Si P_n est régulier, alors $W(z) = 1$ dans (2.16).

Si $r_{m,n} = 0$, alors $\hat{P}_n(z) \equiv P_0(z) \equiv 1$ et $\partial W = \delta = n$ dans (2.16).

Dans ce théorème, les vrais polynômes formels orthogonaux P_n , aussi loin qu'ils existent, sont essentiellement les dénominateurs d'une séquence diagonale $\{r_{n-1,n}\}_{n=0}^\infty$ d'approximants de Padé.

L'existence et l'unicité d'un P_n dépendent seulement de la façon dont la diagonale coupe les blocs de la table (voir figure (2.5)).

Définition 2.15 *Si, pour certains n , il n'existe pas de vrais polynômes orthogonaux formels par rapport à ϕ et si n_j est le plus grand entier plus petit que n pour lequel un polynôme orthogonal formel régulier existe, alors un polynôme $P_n(z) = W(z)P_{n_j}(z)$ avec $W \in P_{n-n_j}$ monique, est appelé un polynôme formel orthogonal déficient normalisé par rapport à ϕ .*

Théorème 2.16 *Etant donnés $f \in \mathcal{P}, n \in \mathbb{N}^+$, un polynôme formel orthogonal P_n est déficient si et seulement si la paire d'indice $(m, n) := (n-1, n)$ est située dans la table de Padé de f , en-dessous de l'antidiagonale d'un bloc carré fini ou d'un bloc nul, i.e. si et seulement si f a un interpolant de Padé (m, n) déficient.*

P_n a alors la forme générale (2.16) où, à nouveau, $\hat{P}_n(z) = P_{n_j}(z) = z^{n-\delta} \hat{q}_{m,n}(z^{-1})$ est monique de degré $(n - \delta)$, $\hat{q}_{m,n}$ est le dénominateur normalisé de l'approximant de Padé (m, n) et $W(z) = z^\delta \omega(z^{-1})$ est un polynôme monique arbitraire de degré exact δ .

Cependant, si P_n est déficient, $q(z) = q_{n-1,n}(z) = z^n P_n(z^{-1})$ n'est pas le second membre

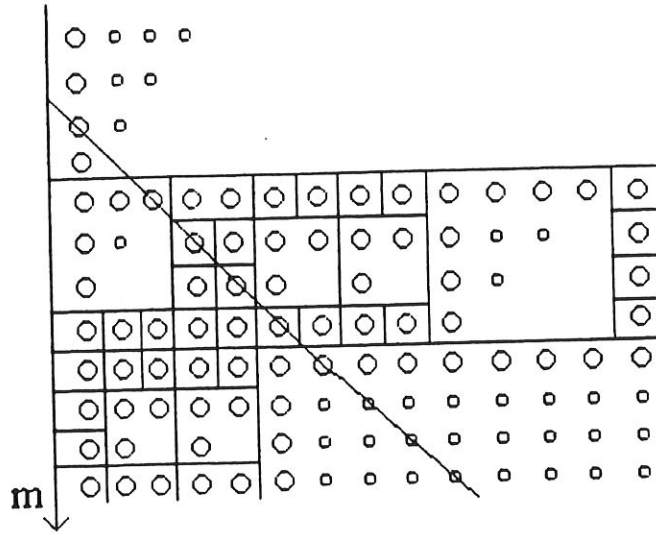


FIG. 2.5 - Le lieu des vrais polynômes orthogonaux formels P_n réguliers (grands cercles) et singuliers (petits cercles) dans la table de Padé de f . Le sommet de la table est composé d'un bloc nul et le bloc inférieur droit, d'un bloc carré. Les "blancs" représentent les entrées des polynômes déficients.

d'une forme de Padé (m,n) de f . Pour obtenir un tel second membre, nous devrions restreindre dans (2.16) le degré de W à $(\delta - \sigma)$, où

$$\sigma = 2n - (n_{j+1} - n_j - 1) \quad (2.17)$$

où n_j (n_{j+1}) est le plus grand (petit) entier au plus (au moins) égal à n ($n+1$) pour lequel un polynôme orthogonal formel régulier existe.

Preuve.

- La première partie du théorème découle du théorème 2.14.
- Si P_n est déficient, nous avons $\sigma \neq 0$ dans (2.12). Il suit donc que (2.16) devient

$$\begin{aligned} P_n(z) &= z^n [z^{-\sigma} \hat{q}_{m,n}(z^{-1}) \underbrace{w(z^{-1})}_{w \in P_{\delta-\sigma}}] \\ &= \underbrace{z^{n-\delta} \hat{q}_{m,n}(z^{-1})}_{\hat{P}_n(z)} \underbrace{z^{\delta-\sigma} w(z^{-1})}_{W(z) \in P_{\delta-\sigma}} \end{aligned}$$

de sorte que le degré de W soit restreint à $(\delta - \sigma)$.

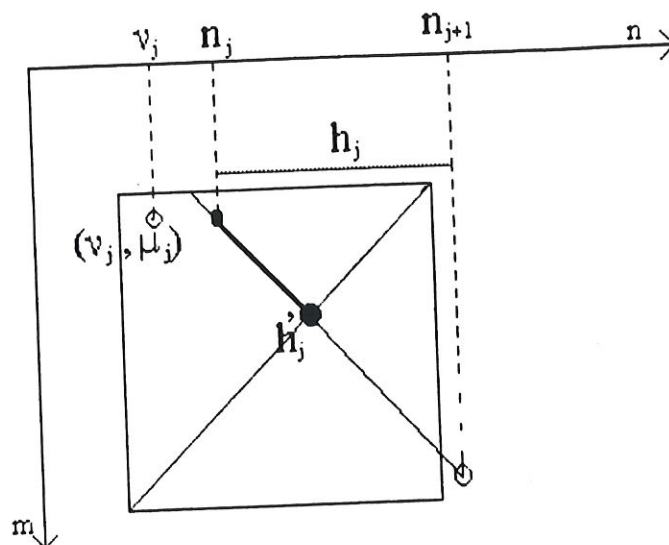


FIG. 2.7 - Lieu des polynômes vrais.

- La relation (2.17) est obtenue de la manière suivante: par (2.15), nous avons $\sigma = m + n - (\mu_j + \nu_j + \Delta_j)$. Par l'égalité $m = n - 1$ et par la figure (2.6), nous pouvons écrire:

$$\sigma = 2n - 1 - \mu_j - \nu_j - \Delta_j = 2n - 1 - n_j + 1 - n_{j+1} + 1 = 2n - (n_j + n_{j+1} - 1) \quad \square$$

Nous avons donc maintenant un ensemble complet de polynômes orthogonaux formels et le théorème suivant est une reformulation utile des théorèmes (2.14) et (2.16).

Théorème 2.17 Soit $0 = n_0 < n_1 < n_2 < \dots$ les indices pour lesquels un polynôme orthogonal formel régulier P_n existe. Alors, pour tout $n \in \mathbb{N}$, les polynômes orthogonaux formels P_n ont la forme

$$P_n(z) := W_{n-n_j}(z)P_{n_j}(z) \quad \text{si } n_j \leq n < n_{j+1} \quad (2.18)$$

avec W_{n-n_j} un polynôme monique, arbitraire de degré exact $\delta = n - n_j$. Si nous posons $h_j := n_{j+1} - n_j$, $h'_j := \lfloor \frac{h_j-1}{2} \rfloor$ on a pour $n_j \leq n < n_{j+1}$, $j \in \mathbb{N}$:

- (i) P_n est un vrai polynôme orthogonal formel si et seulement si $n_j \leq n \leq n_j + h'_j$ (voir figure (2.7)).
- (ii) P_n est un polynôme orthogonal formel régulier si et seulement si $n = n_j$ (voir figure (2.8)).

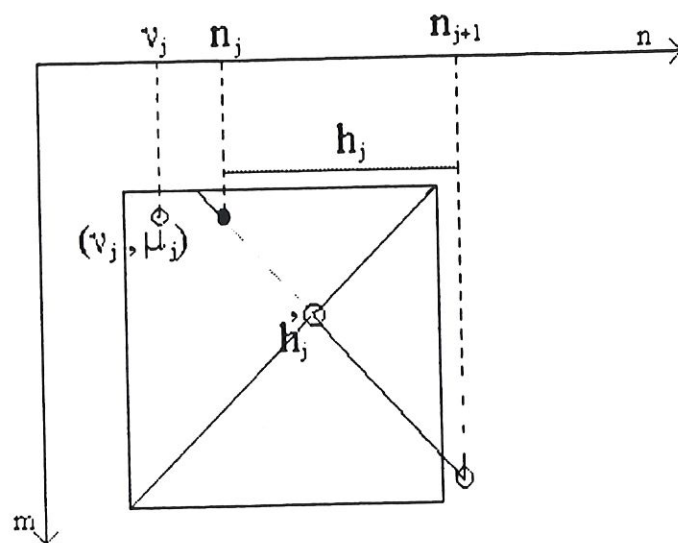


FIG. 2.8 - Lieu des polynômes réguliers.

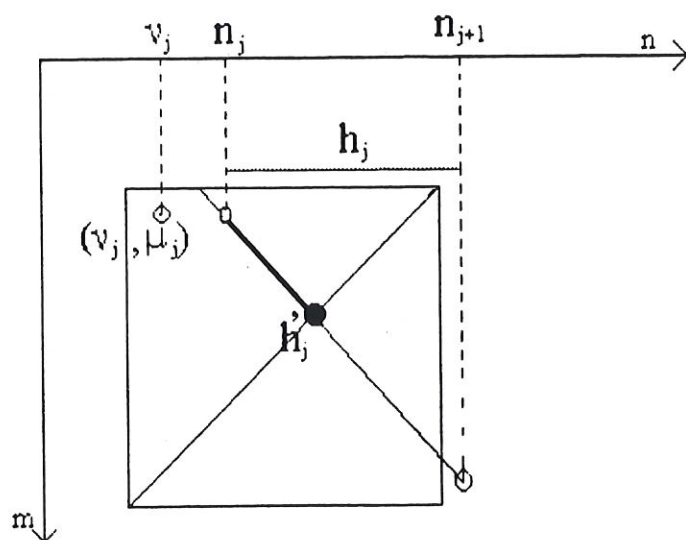


FIG. 2.9 - Lieu des polynômes singuliers.

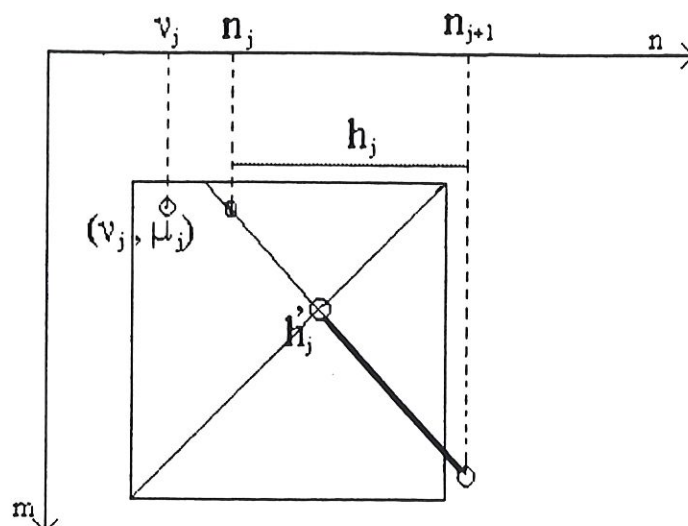


FIG. 2.10 - Lieu des polynômes déficients.

(iii) P_n est un polynôme orthogonal formel singulier si et seulement si $n_j < n \leq n_j + h_j'$ (voir figure (2.9)).

(iv) P_n est un polynôme orthogonal formel déficient si et seulement si $n_j + h_j' < n < n_{j+1}$ voir figure (2.10)).

où $[x]$ représente le plus grand entier non supérieur à x .

Soulignons que si f est rationnelle, il y a un nombre fini de polynômes orthogonaux formels P_n , réguliers, $j = 0, \dots, J$.

Chapitre 3

Les formules d'orthogonalité et de récurrence des polynômes orthogonaux formels.

Dans le chapitre précédent, pour chaque n , nous avons défini des polynômes formels normalisés P_n orthogonaux par rapport à la fonctionnelle ϕ . A présent, nous examinons les différentes propriétés d'orthogonalité de la séquence $\{P_n\}$ pour en déduire une formule de récurrence. Par définition, les polynômes orthogonaux formels sont orthogonaux à tous les polynômes de degré inférieur. Ils peuvent cependant être orthogonaux à eux-mêmes, ce qui se passe quand un polynôme orthogonal formel correspond à un approximant de Padé se situant dans la table de Padé de f au-dessus de l'antidiagonale de ce bloc. D'autre part, les polynômes orthogonaux formels déficients se situant en-dessous de l'antidiagonale ne sont pas orthogonaux à tous les polynômes de degré inférieur. Ces propriétés sont l'objet du théorème suivant.

Théorème 3.1 *Soit $\{W_m\}_{m=0}^\infty$, une séquence de polynômes moniques $W_m \in \mathcal{P}_m$. Pour un polynôme orthogonal formel P_n avec $n_j \leq n < n_{j+1}$, nous avons :*

(i) *Si $j < J$ et, soit $j > 0$, soit $f(z) \neq \mathcal{O}_+(z)$, i.e. si l'entrée $(n_j - 1, n_j)$ de la table de Padé de f se situe sur le bord supérieur ou à gauche d'un bloc carré fini, alors*

$$\begin{aligned} \phi(pP_n) &= 0 & (\forall p \in P_{\hat{n}-1}) \\ \phi(z^{\hat{n}}P_n) &\neq 0 \end{aligned} \tag{3.1}$$

où

$$\hat{n} := 2n_j - n + h_j - 1 = n_j + n_{j+1} - n - 1. \tag{3.2}$$

et, le dernier terme de (2.7) devient $\mathcal{O}_+(\hat{n})$. Nous avons donc que $f(z) = \mathcal{O}_+(z^k)$ et, par (2.7), que $f(z) = \mathcal{O}_+(\hat{n})$. Ceci entraîne

$$\hat{n} = k = n_1 - 1.$$

Finalement, si n augmente de n_0 à $n_1 - 1$, par un raisonnement analogue à celui effectué au point (i) de ce théorème, nous obtenons

$$\hat{n} = n_1 - n - 1. \quad \square$$

Les résultats suivants peuvent être déduits du théorème 3.1 :

Corollaire 3.2 *Sous les hypothèses de la partie (i) du théorème (3.1) et en prenant comme convention $n_i \leq n' < n_{i+1}$, $n_j \leq n < n_{j+1}$, nous avons :*

$$\phi(P_{n'}P_n) = 0 \begin{cases} \text{si } i \neq j \\ \text{ou } i = j \text{ et } n' + n < n_j + n_{j+1} - 1 \end{cases} \quad (3.6)$$

et, pour certains δ_j non nuls indépendants de $n - n_j$ et $n' - n_j$,

$$\phi(P_{n'}P_n) := \delta_j \neq 0 \text{ si } i = j \text{ et } n' + n = n_j + n_{j+1} - 1.$$

Si $i = j$ et $n' + n \geq n_j + n_{j+1}$, la valeur de $\phi_0(P_{n'}P_n)$ ne peut être prédite.

Preuve. (3.6) est une conséquence immédiate de (3.1). En effet, si $i \neq j$, alors soit $n' < n_j \leq n$, soit $n < n_i \leq n'$; si $i = j$, $\phi(P_{n'}P_n) = 0$ si $n' \leq \hat{n}$ où \hat{n} est défini en (3.2). Semblablement, $\phi(P_{n'}P_n) = \delta_j \neq 0$ est une conséquence de la deuxième équation de (3.1) et de l'équation (3.2). δ_j dépend uniquement de j car tout P_n est monique et ϕ est linéaire :

$$\delta_j = \phi(z^{n_{j+1}-1}P_{n_j}) = \phi(z^{h_j-1}P_{n_j}^2). \quad \square$$

Corollaire 3.3 (i) *Un polynôme orthogonal formel P_n est orthogonal à tous les polynômes de degré inférieur et non à lui-même si et seulement si l'entrée correspondante $(n_j - 1, n_j)$ se situe, dans la table de Padé de f , sur l'antidiagonale d'un bloc fini ou nul.*

(ii) *P_n est orthogonal à tous les polynômes de degré inférieur ou égal si et seulement si l'entrée $(n_j - 1, n_j)$ se situe au-dessus de l'antidiagonale d'un bloc fini ou nul ou encore, d'un bloc infini non nul.*

(iii) *P_n n'est pas orthogonal à tous les polynômes de degré inférieur si et seulement si l'entrée $(n_j - 1, n_j)$ se situe en-dessous de l'antidiagonale d'un bloc fini ou nul.*

Corollaire 3.4 *Pour tout $n \in \mathbb{N}$, il existe un polynôme de degré au plus n qui est, par rapport à ϕ , orthogonal à tout polynôme de degré inférieur à n .*

A présent, intéressons-nous aux formules de récurrence. Par la définition (2.18), tous les polynômes P_n avec $n_j < n < n_{j+1}$ sont générés à partir des polynômes orthogonaux formels P_n réguliers. De plus, ils peuvent être calculés récursivement. Par exemple, si $W_m(z) = z^m$, alors :

$$P_{n+1}(z) = zP_n(z), \quad (n_j \leq n \leq n_{j+1} - 2).$$

Tandis que, si les polynômes W_m sont choisis pour satisfaire une récurrence à trois termes :

$$W_{m+1}(z) = (z - \alpha_m^W)W_m(z) - \beta_m^W W_{m-1}(z) \quad (m \in \mathbb{N}) \quad (3.7)$$

(avec $W_0(z) := 1, W_{-1}(z) := 0, \beta_0^W := 0$), alors, on a

$$P_{n+1}(z) = (z - \alpha_{n-n_j}^W)P_n(z) - \beta_{n-n_j}^W P_{n-1}(z) \quad (n_j \leq n \leq n_{j+1} - 2). \quad (3.8)$$

Une fois les relations d'orthogonalité de (3.1) établies, la formule de récurrence à trois termes pour $\{P_{n_j}\}_{j=0}^J (J \leq \infty)$ s'écrit facilement.

Etant donné que ces polynômes sont moniques de degré n_j , ils satisfont à la récurrence :

$$P_{n_j+1}(z) = (W_{h_j}(z) - t_{j,j}(z))P_{n_j}(z) - t_{j-1,j}(z)P_{n_j-1}(z) - \dots - t_{0,j}P_0(z)$$

avec $t_{i,j} \in \mathcal{P}_{h_j-1}$, $i = 0, \dots, j$. Multiplions cette relation par z^{n_i+k} ($k = 0, \dots, h_j - 1$; $i = 0, \dots, j$) et appliquons la fonctionnelle ϕ .

Puisque $n_i + k \leq n_{i+1} - 1$, le membre de gauche devient nul et nous obtenons, pour chaque paire (k, i) une équation linéaire :

$$\sum_{s=0}^j \phi(z^{n_i+k} t_{s,j} P_{n_s}) = \phi(z^{n_i+k} W_{h_j} P_{n_j}) \quad (k = 0, \dots, h_j - 1; \quad i = 0, \dots, j).$$

En vue de satisfaire à (3.1), (3.2) et (3.3), nous avons pour le membre de droite :

$$\phi(z^{n_i+k} W_{h_j} P_{n_j}) = \begin{cases} 0 & \text{si } i < j - 1 \\ 0 & \text{si } i = j - 1 \text{ et } k \leq h_{j-1} - 2 \\ \neq 0 & \text{si } i = j - 1 \text{ et } k = h_{j-1} - 1 \end{cases}$$

Pour le membre de gauche,

$$\phi(z^{n_i+k} t_{s,j} P_{n_s}) = \begin{cases} 0 & \text{si } i < s \\ 0 & \text{si } i = s \text{ et } k + \partial t_{s,j} \leq h_s - 2 \\ \neq 0 & \text{si } i = s \text{ et } k + \partial t_{s,j} = h_s - 1 \end{cases}$$

Le système se réduit donc à une paire couplée de systèmes, l'un homogène :

$$\sum_{s=0}^i \phi(z^{n_i+k} t_{s,j} P_{n_s}) = 0 \quad (3.9)$$

où $(k = 0, \dots, h_i - 1; i = 0, \dots, j - 2 \quad \text{et} \quad k = 0, \dots, h_{j-1} - 2; i = j - 1)$

constitué de

$$\sum_{i=0}^{j-1} h_i - 1 = n_j - 1$$

équations pour j polynômes $\{t_{s,j}\}_{s=0}^{j-1}$ avec un degré total de liberté n_j .
L'autre non-homogène :

$$\sum_{s=0}^i \phi(z^{n_i+k} t_{s,j} P_{n_s}) = \phi(z^{n_i+k} W_{h_j} P_{n_j}) \quad (3.10)$$

où $(k = 0, \dots, h_{j-1}; i = j \quad \text{et} \quad k = h_{j-1} - 1, i = j - 1)$

constitué de $h_j + 1 = n_{j+1} - n_j + 1$ équations et contenant en plus les polynômes $t_{j,j}$ avec h_j paramètres libres. Le système en entier, tout comme sa partie homogène, est triangulaire.

Propriété 3.5 *Toutes les inconnues apparaissant dans le système homogène sont nulles sauf une.*

Preuve. Procédons par récurrence. Supposons $t_{0,j} \neq 0$. Choisissons k tel que $k + \partial t_{0,j} = h_0 - 1$. On a $\phi(z^{n_0+k} t_{0,j} P_{n_0}) \neq 0$. Mais cela est en contradiction avec (3.1) et (3.2). Par conséquent, $t_{0,j} = 0$. Supposons maintenant $t_{0,j} = t_{1,j} = \dots = t_{i-1,j} \neq 0$ ($i \leq j - 1$). Choisissons k tel que $k + \partial t_{i,j} = h_i - 1$. Par un raisonnement analogue au précédent, nous concluons que $t_{i,j} = 0$ excepté dans le cas où $i = j - 1$ où il n'y a pas d'équation pour $k = h_{j-1} - 1$ telle que $t_{j-1,j}$ pourrait être un polynôme non nul de degré 0, i.e., une constante.

Inversément, supposons que

$$t_{0,j}(z) = t_{1,j}(z) = \dots = t_{j-2,j}(z) = 0 \quad t_{j-1,j}(z) \in \mathbb{C}$$

est une solution du système homogène (3.9), et ce doit être aussi la solution générale du système complet. En insérant cette solution dans (3.10) et en remplaçant $t_{j-1,j}$ et $t_{j,j}$ par β_j (une constante) et a_j (un polynôme), respectivement, (3.10) est finalement réduit à :

$$\beta_j \phi(z^{n_j-1} P_{n_{j-1}}) = \phi(z^{n_j-1} W_{h_j} P_{n_j}), \quad (3.11)$$

$$\phi(z^{n_j+k} a_j P_{n_j}) + \beta_j \phi(z^{n_j+k} P_{n_{j-1}}) = \phi(z^{n_j+k} W_{h_j} P_{n_j}) \quad (3.12)$$

où $k = 0, \dots, h_j - 1$

i.e., une équation linéaire individuelle pour la constante β_j et un système de h_j équations linéaires pour les coefficients h_j de a_j .

Rappelons que (3.11) et (3.12) signifient juste que

$$\begin{aligned} \beta_j P_{n_j-1} - W_{h_j} P_{n_j} &\perp z^{n_j-1}, \\ a_j P_{n_j} + \beta_j P_{n_j-1} - W_{h_j} P_{n_j} &\perp z^{n_j+k} \end{aligned} \quad (k = 0, \dots, h_j - 1) \quad (3.21)$$

où on note " \perp " la "pseudo-orthogonalité" par rapport au produit interne bilinéaire indéfini ϕ :

$$(p, q) \longrightarrow \phi(p, q).$$

Grâce à (3.1) et (3.2), les membres de gauche de (3.21) sont orthogonaux à un polynôme $p \in \mathcal{P}_{n_j-2}$. Par conséquent, dans le membre de droite, les monômes z^{n_j-1} et z^{n_j+k} ($k=0, \dots, h_j-1$) peuvent être remplacés par d'autres polynômes de même degré. Un choix naturel est de les remplacer par P_{n_j-1} et P_{n_j+k} ($k=0, \dots, h_j-1$), respectivement. Cela implique que les termes en β_j dans la deuxième équation de (3.21) disparaissent. On obtient :

$$\beta_j \phi(P_{n_j-1} P_{n_j-1}) = \phi(P_{n_j-1} W_{h_j} P_{n_j}), \quad (3.22)$$

$$\phi(a_j P_{n_j+k} P_{n_j}) = \phi(P_{n_j+k} W_{h_j} P_{n_j}) \quad (k = 0, \dots, h_j - 1). \quad (3.23)$$

De plus, puisque $W_{h_j} P_{n_j}$ est une combinaison linéaire de $z P_{n_{j+1}-1}, P_{n_{j+1}-1}, P_{n_{j+1}-2}, \dots, P_{n_j}$ par (3.18) et que tous ces polynômes, excepté le premier, sont orthogonaux à P_{n_j-1} , on peut remplacer le membre de droite de (3.22) par $\phi(z P_{n_j-1} P_{n_{j+1}-1})$. Dans (3.23), cette simplification ne s'applique pas mais on peut substituer $W_{h_j} P_{n_j}$ selon (3.18) si (3.7) est respectée.

Théorème 3.9 *Si (3.7) est respectée, le système linéaire (3.11) et (3.12) pour calculer les polynômes a_j et les constantes β_j pour la récurrence (3.13) peut être remplacé par le système équivalent :*

$$\begin{aligned} \beta_j \phi(P_{n_j-1} P_{n_j-1}) &= \phi(z P_{n_j-1} P_{n_{j+1}-1}), \\ \phi(a_j P_{n_j+k} P_{n_j}) &= \phi(z P_{n_j+k} P_{n_{j+1}-1}) - \alpha_{h_j-1}^W \phi(P_{n_j+k} P_{n_{j+1}-1}) \\ &\quad - \beta_{h_j}^W \phi(P_{n_j+k} P_{n_{j+1}-2}) \end{aligned} \quad (k = 0, \dots, h_j - 1) \quad (3.24)$$

se composant d'une équation simple pour β_j et, si a_j est exprimé comme en (3.19) d'un système triangulaire inférieure droit pour les coefficients $\alpha_{j,s}$ de a_j . Si (3.7) n'est pas satisfaite, le système (3.24) n'est pas valide.

Chapitre 4

Interprétation matricielle.

Le but de ce chapitre est de réécrire, sous forme matricielle, les différentes formules de récurrence et d'orthogonalité des polynômes vues au chapitre 3. Cette interprétation matricielle nous permettra, dans le dernier chapitre, de nous intéresser d'un peu plus près aux problèmes des "breakdowns" lors de l'application de l'algorithme de biorthogonalisation.

Soit w_m , des vecteurs lignes

$$w_m := [W_0, W_1, \dots, W_m] \quad (m \in \mathbb{N})$$

et T_m^W , des matrices tridiagonales :

$$T_m^W := \begin{pmatrix} \alpha_0^W & \beta_0^W & & & \\ 1 & \alpha_1^W & \beta_1^W & & \\ & 1 & \ddots & \ddots & \\ & & \ddots & 1 & \beta_m^W \\ & & & & \alpha_m^W \end{pmatrix} \quad (m \in \mathbb{N}).$$

Grâce à ces nouveaux éléments, nous pouvons reformuler la récurrence à 3 termes (3.7)

$$zw_m(z) = w_m(z)T_m^W + [0, \dots, 0, W_{m+1}(z)] \quad (m \in \mathbb{N}). \quad (4.1)$$

On remarque que chaque racine de W_{m+1} est une valeur propre de T_m^W .

Propriété 4.1 *Si W_{m+1} admet une racine de multiplicité $\mu > 1$, la différentiation de (4.1) montre que les vecteurs $w_m'(z), \dots, w_m^{(\mu-1)}(z)$ sont des vecteurs principaux. Ceci entraîne que T_m^W est non dégénérée, i.e., que toutes ses valeurs propres ont une multiplicité géométrique d'une unité [9].*

Avant de démontrer cette propriété, définissons la notion de "vecteur principal" [6].

Définition 4.2 *Un vecteur satisfaisant à $(A - \lambda_i I)^j x_k = 0$ sans satisfaire à une relation $(A - \lambda_i I)^i, i < j$, est appelé vecteur principal de degré j correspondant à λ_j .*

En particulier, un vecteur propre est un vecteur principal de degré 1.

Preuve (Propriété (4.1)). Soit α , une racine de multiplicité μ de $W_{m+1}(z)$:

$$\alpha w_m(\alpha) = w_m(\alpha) T_m^W$$

alors

$$w_m(\alpha)(\alpha I - T_m^W) = 0.$$

Après dérivation, nous obtenons

$$\begin{aligned} w_m(\alpha) + \alpha w_m'(\alpha) &= w_m'(\alpha) T_m^W, \\ w_m'(\alpha)(\alpha I - T_m^W) &= -w_m(\alpha), \\ w_m'(\alpha)(\alpha I - T_m^W)^2 &= 0. \end{aligned}$$

De proche en proche, nous avons

$$w_m^{(k)}(\alpha)(\alpha I - T_m^W)^{k+1} = 0 \quad (0 < k \leq \mu - 1).$$

$w_m(\alpha)$ est donc un vecteur propre et, par la définition (4.2), $w_m^{(k)}(\alpha)$ sont des vecteurs principaux de degré k ($0 < k \leq \mu - 1$). On voit que la multiplicité de la racine α correspond au nombre μ de vecteurs principaux (le premier étant vecteur propre). La matrice T_m^W ne comprend alors qu'un seul bloc de Jordan associé à la racine α . Comme la multiplicité géométrique de la racine α est égale au nombre de blocs de Jordan associés à cette même racine α , nous en déduisons que la multiplicité géométrique est égale à 1. Ceci entraîne la dégénérescence de la matrice T_m^W . \square

Il est important que la relation (4.1) reste vérifiée pour une récurrence incluant davantage d'itérés précédents et pour une matrice T_m^W d'Hessenberg supérieure et non plus tridiagonale.

Soit n , les indices pour lesquels il n'existe pas de polynômes formels orthogonaux réguliers. Supposons que P_n est défini par (2.18) et que la séquence $\{W_m\}$ satisfait à la récurrence à trois termes (3.7). Etant donné que (3.7) entraîne (3.8), valide pour $n_j \leq n \leq n_{j+1} - 2$, nous obtenons :

$$z[P_{n_j}(z), \dots, P_{n_{j+1}-2}(z)] = [P_{n_j}(z), \dots, P_{n_{j+1}-2}(z)]T_{h,-2}^W + [0, \dots, 0, P_{n_{j+1}-1}(z)].$$

et

$$J := \begin{pmatrix} 0 & & & \\ 1 & 0 & & \\ & 1 & 0 & \\ & & \ddots & \ddots \end{pmatrix}.$$

Soit la matrice triangulaire supérieure unitaire infinie :

$$P := \begin{pmatrix} 1 & \pi_{0,1} & \pi_{0,2} & & \\ & 1 & \pi_{1,2} & & \\ & & 1 & \ddots & \\ & & & \ddots & \\ & & & & \ddots \end{pmatrix} \quad (4.5)$$

dont les éléments de la $(n+1)$ -ième colonne sont les coefficients de P_n (cfr (2.4)). Les éléments diagonaux sont tous égaux à 1 puisque les polynômes P_n sont moniques. En utilisant ces notations, nous pouvons dès lors écrire :

$$zZ(z) = Z(z)J$$

et

$$p(z) = Z(z)P, \quad zp(z) = Z(z)JP$$

tels que (4.4) devienne :

$$Z(z)JP = Z(z)PH. \quad (4.6)$$

Chaque composante des vecteurs lignes de la relation ci-dessus constitue un polynôme. Ceci entraîne que les matrices contenant les coefficients doivent être égales :

$$JP = PH.$$

En nous intéressant aux sous-matrices principales, nous obtenons la relation correspondant à (4.4). En résumé, on obtient :

Corollaire 4.5 *Sous les hypothèses du théorème (4.3), nous avons*

$$JP = PH.$$

De plus, si J_n , P_n et H_n sont les sous-matrices principales d'ordre $(n+1)$ de J , P et H , alors

$$J_n P_n = P_n H_n + \begin{pmatrix} \pi_{0,n+1} \\ \pi_{1,n+1} \\ \dots \\ \pi_{n,n+1} \end{pmatrix} (0, \dots, 0, 1).$$

Une autre interprétation matricielle importante est immédiatement obtenue à partir du corollaire (3.2).

Théorème 4.6 Soit $p := [P_0, P_1, \dots]$ défini comme dans le théorème (4.3) et soit $\phi(p^T p)$, la matrice infinie obtenue par application de ϕ à la matrice infinie de rang 1, $(p^T p)$. Alors

$$\phi(p^T p) = D$$

où D est une matrice diagonale constituée de blocs,

$$D := \text{bloc diag}[D_0, D_1, \dots] \quad (4.7)$$

où les D_j sont des blocs carrés de taille h_j ($j = 0, 1, \dots, J$). Les sous-matrices D_j sont triangulaires inférieures droites symétriques de la forme

$$D_j := \begin{pmatrix} & & & \delta_j \\ & & & * \\ & & \delta_j & * \\ & \dots & \dots & \vdots \\ & \delta_j & \dots & * \\ \delta_j & * & \dots & * & * \end{pmatrix}.$$

Nous avons $\delta_j := \phi(P_{n_j} P_{n_{j+1}-1}) \neq 0$. Les étoiles indiquent les éléments qui peuvent être non nuls dans la matrice D_j . Si $W_m(z) = z^m$, D_j est une matrice de Hankel. Si $J < \infty$, D_J est une matrice nulle infinie et $\delta_J = 0$. De plus, prenons P la matrice unitaire triangulaire de (4.5) et posons M la matrice infinie des moments de la fonctionnelle ϕ ,

$$M := \begin{pmatrix} \psi_0 & \psi_1 & \psi_2 & \dots \\ \psi_1 & \psi_2 & & \\ \psi_2 & & & \\ \vdots & & & \end{pmatrix}.$$

Nous avons alors

$$P^T M P = D.$$

Par conséquent, si $R := P^{-1}$ est une matrice unitaire triangulaire supérieure dont les colonnes sont constituées des coefficients des monômes, exprimés comme combinaison linéaire des polynômes formels orthogonaux, alors

$$M = R^T D R.$$

Cette dernière équation nous donne une décomposition LDU en blocs symétriques de M .

Chapitre 5

Les polynômes orthogonaux et l'algorithme de biorthogonalisation de Lanczos.

Dans ce chapitre, nous exprimerons les séquences de vecteurs $\{x_n\}$ et $\{y_n\}$ à partir de séquences de polynômes. Ensuite, grâce aux différentes expressions d'orthogonalité et de récurrence des polynômes vues au chapitre 3, nous pourrions formuler l'algorithme BO non générique.

5.1 L'orthogonalité et l'algorithme de Lanczos.

Soit K_n l'espace de Krylov où $x, y \in K_n$. On peut représenter

$$\begin{aligned} x &= P_n(A)x_0, \\ y &= \bar{P}_n(\tilde{A})y_0 \end{aligned} \tag{5.1}$$

où P_n et \bar{P}_n sont des polynômes de degré n dont les arguments sont A et \tilde{A} respectivement. Les coefficients de $\bar{P}_n(\tilde{A})$ sont les complexes conjugués des coefficients de $P_n(A)$.

Associée à la séquence $\{x_n\}_{n=0}^\nu$ définie par l'algorithme BO, nous avons une séquence $\{P_n\}_{n=0}^\nu$ de polynômes et, grâce à (1.11), nous avons une formule de récurrence à trois termes pour ces mêmes polynômes (de même pour les polynômes associés à la séquence $\{y_n\}_{n=0}^\nu$).

En réécrivant la propriété (1.2) avec les éléments de (5.1), on obtient :

Nous rappelons que $\{n_j\}_{j=0}^J$ représente la séquence d'indices pour lesquels le polynôme P_n est régulier (i.e. unique).

Par (5.3), (5.5) et la linéarité de ϕ , nous obtenons maintenant une règle générale :

$$\begin{aligned}\langle y_m, A^k x_n \rangle &= \overline{\Gamma_m} \Gamma_n \langle y_0, A^k P_m(A) P_n(A) x_0 \rangle \\ &= \overline{\Gamma_m} \Gamma_n \phi_0(\xi^k P_m P_n)\end{aligned}$$

Après avoir établi ce lien avec les polynômes orthogonaux formels, nous pouvons aisément tirer des conclusions à partir du chapitre 3, en particulier, un résultat fondamental concernant l'orthogonalité des polynômes :

Théorème 5.2 *Soit A , une matrice carrée finie et x_0, y_0 , deux vecteurs initiaux non-orthogonaux. Définissons deux séquences $\{x_n\}_{n=0}^\infty, \{y_n\}_{n=0}^\infty$ par (5.5), où P_n est un n -ième polynôme orthogonal pour la fonctionnelle ϕ , et $\Gamma_n, \overline{\Gamma}_n$ les facteurs scalaires arbitraires non nuls.*

Soit $\{n_j\}_{j=0}^J$ ($J \leq \infty$) les séquences d'indices pour lesquels P_n est régulier. Alors, pour $n < n_J$ (ou pour tout n si $J = \infty$),

$$\begin{aligned}x_n &\perp L_{\hat{n}} \quad , \quad y_n \perp K_{\hat{n}}, \\ \langle y_n, x_{\hat{n}} \rangle &= \overline{\Gamma_n} \Gamma_{\hat{n}} \delta_j \neq 0, \\ \langle y_{\hat{n}}, x_n \rangle &= \overline{\Gamma_{\hat{n}}} \Gamma_n \delta_j \neq 0\end{aligned}$$

où n et \hat{n} sont liés par (5.6) et où $\delta_j := \phi(P_{\hat{n}} P_n)$.

Si $J < \infty$, alors

$$x_n \perp L_m, \quad y_n \perp K_m \quad (n \geq n_J, \quad m \in \mathbb{N}^+) \quad (5.7)$$

Notons que nous ne rejetons ni le cas où $x_n = 0$, ni le cas où $y_n = 0$ pour $n \geq n_J$. D'autre part, il peut arriver que (5.7) soit valable bien que $x_n \neq 0$ et $y_n \neq 0$.

5.2 Les formules de récurrence et l'algorithme de Lanczos .

A présent, notre objectif est d'écrire, à partir des formules de récurrence pour P_n données dans le chapitre 3, un algorithme récursif pour calculer les séquences $\{x_n\}$ et $\{y_n\}$. Nous supposons que, pour les degrés n où le polynôme orthogonal formel n'est pas unique, P_n est choisi comme en (2.18) où $\{W_m\}$ est une base monique fixée satisfaisant à

la récurrence à trois termes (3.7) de telle manière que (3.8) soit vérifiée. Cette récurrence transforme (5.5) en :

$$x_{n+1} = (Ax_n - x_n \alpha_{n-n_j}^W) \gamma_{n+1,1} - x_{n-1} \beta_{n-n_j}^W \gamma_{n+1,2} \quad (5.8)$$

où $n_j \leq n \leq n_{j+1} - 2$

$$y_{n+1} = (\tilde{A}y_n - y_n \overline{\alpha_{n-n_j}^W}) \bar{\gamma}_{n+1,1} - y_{n-1} \overline{\beta_{n-n_j}^W} \bar{\gamma}_{n+1,2}$$

où $n_j \leq n \leq n_{j+1} - 2$ et

$$\gamma_{n,i} := \frac{\Gamma_n}{\Gamma_{n-i}}, \quad \bar{\gamma}_{n,i} := \frac{\bar{\Gamma}_n}{\bar{\Gamma}_{n-i}} \quad (n \in \mathbb{N}, i \in \mathbb{N})$$

tels que, en vue de satisfaire à $\Gamma_0 = \bar{\Gamma}_0 = 1$

$$\begin{aligned} \Gamma_n &= \gamma_{1,1} \gamma_{2,1} \dots \gamma_{n,1} & \gamma_{n,i} &= \gamma_{n,1} \gamma_{n-1,1} \dots \gamma_{n-i+1,1}, \\ \bar{\Gamma}_n &= \bar{\gamma}_{1,1} \bar{\gamma}_{2,1} \dots \bar{\gamma}_{n,1} & \bar{\gamma}_{n,i} &= \bar{\gamma}_{n,1} \bar{\gamma}_{n-1,1} \dots \bar{\gamma}_{n-i+1,1}. \end{aligned} \quad (5.9)$$

Dans le chapitre 3, (3.17) donnait une formule de récurrence :

$$\begin{aligned} P_{n_{j+1}}(z) &= z P_{n_{j+1}-1}(z) - (\alpha_{h_j-1,j} + \alpha_{h_j-1}^W) P_{n_{j+1}-1}(z) - (\alpha_{h_j-2,j} + \beta_{h_j-1}^W) P_{n_{j+1}-2}(z) \\ &\quad - \alpha_{h_j-3,j} P_{n_{j+1}-3}(z) - \dots - \alpha_{0,j} P_{n_j}(z) - \beta_j P_{n_{j-1}}(z), \\ j &= 0, \dots, J-1. \end{aligned}$$

Si nous remplaçons z par A ou \tilde{A} dans cette formule et tenons compte de (5.5), nous obtenons :

$$\begin{aligned} x_{n_{j+1}} &= [Ax_{n_{j+1}-1} - x_{n_{j+1}-1}(\alpha_{h_j-1,j} + \alpha_{h_j-1}^W)] \gamma_{n_{j+1},1} \\ &\quad - x_{n_{j+1}-2}(\alpha_{h_j-2,j} + \beta_{h_j-1}^W) \gamma_{n_{j+1},2} - x_{n_{j+1}-3} \alpha_{h_j-3,j} \gamma_{n_{j+1},3} - \dots \\ &\quad - x_{n_j} \alpha_{0,j} \gamma_{n_{j+1},h_j} - x_{n_{j-1}} \beta_j \gamma_{n_{j+1},h_j+h_{j-1}}, \\ y_{n_{j+1}} &= [\tilde{A}y_{n_{j+1}-1} - y_{n_{j+1}-1}(\alpha_{h_j-1,j} + \alpha_{h_j-1}^W)] \bar{\gamma}_{n_{j+1},1} \\ &\quad - y_{n_{j+1}-2}(\alpha_{h_j-2,j} + \beta_{h_j-1}^W) \bar{\gamma}_{n_{j+1},2} - y_{n_{j+1}-3} \overline{\alpha_{h_j-3,j}} \bar{\gamma}_{n_{j+1},3} - \dots \\ &\quad - y_{n_j} \overline{\alpha_{0,j}} \bar{\gamma}_{n_{j+1},h_j} - y_{n_{j-1}} \overline{\beta_j} \bar{\gamma}_{n_{j+1},h_j+h_{j-1}}. \end{aligned} \quad (5.10)$$

Dans le cas où $h_j = 1$, ces deux formules se simplifient :

$$\begin{aligned} x_{n_{j+1}} &= [Ax_n - x_n(\alpha_{0,j} + \alpha_0^W)] \gamma_{n_{j+1},1} - x_{n_{j-1}} \beta_j \gamma_{n_{j+1},h_{j-1}+1}, \\ y_{n_{j+1}} &= [\tilde{A}y_n - y_n(\alpha_{0,j} + \alpha_0^W)] \bar{\gamma}_{n_{j+1},1} - y_{n_{j-1}} \overline{\beta_j} \bar{\gamma}_{n_{j+1},h_{j-1}+1}. \end{aligned} \quad (5.11)$$

Les séquences $\{x_j\}$ et $\{h_j\}$ ne sont bien évidemment pas connues à l'avance. Grâce à (3.4), nous rappelons que $h_j := n_{j+1} - n_j$ peut être déterminé selon

$$h_j = 1 + \min\{k \in \mathbb{N}; \langle y_{n_j+k}, x_{n_j} \rangle \neq 0\}. \quad (5.12)$$

Notons que $h_0 = 1$ puisque nous avons supposé que $\langle y_0, x_0 \rangle \neq 0$.

Pour déterminer les coefficients en (5.10), nous appliquerons le théorème 3.9 qui donne :

$$\beta_j \gamma_{n_{j+1}-1, h_j+h_{j-1}-1} \langle y_{n_{j-1}}, x_{n_{j-1}} \rangle = \langle y_{n_{j-1}}, Ax_{n_{j+1}-1} \rangle, \quad (5.13)$$

$$\begin{aligned} & \sum_{s=1}^{k+1} \alpha_{h_j-s, j} \gamma_{n_{j+1}-1, s-1} \langle x_{n_{j+1}-s}, y_{n_{j+1}-k} \rangle \\ &= \langle y_{n_{j+1}-k}, A(x_{n_{j+1}-1} - x_{n_{j+1}-1} \alpha_{h_j-1}^W - x_{n_{j+1}-2} \beta_{h_j-1}^W \gamma_{n_{j+1}-1, 1}) \rangle, \end{aligned} \quad (5.14)$$

$$k = 0, 1, \dots, h_j - 1.$$

Comme dans le théorème 3.9, il s'agit d'une équation linéaire pour la constante β_j et d'un système triangulaire inférieur pour les coefficients $\alpha_{s,j}$ des polynômes $a_j(z)$. Si $h_j = 1$, le dernier système se réécrit sous forme d'une seule équation de telle manière qu'on ait les deux formules explicites :

$$\begin{aligned} \beta_j &:= \frac{1}{\gamma_{n_j, h_j-1}} \frac{\langle y_{n_j-1}, Ax_{n_j} \rangle}{\langle y_{n_j-1}, x_{n_j-1} \rangle}, \\ \alpha_{0,j} &:= \frac{\langle y_{n_j}, Ax_{n_j} - x_{n_j} \alpha_0^W \rangle}{\langle y_{n_j}, x_{n_j} \rangle}. \end{aligned}$$

Ce sont les formules de l'algorithme BO classique. Celles-ci complètent l'algorithme BO vu au chapitre 1.

5.3 Algorithme non générique de biorthogonalisation.

Les séquences de vecteurs $\{x_n\}_{n=0}^\infty$ et $\{y_n\}_{n=0}^\infty$ du théorème (5.2) peuvent être construites par un processus récurrent qui, pour $j = 0, 1, \dots, J$ ($J \leq \infty$) consiste en :

- $\{x_{n_j+k}\}_{k=1}^{h_j-1}$, $\{y_{n_j+k}\}_{k=1}^{h_j-1}$ et h_j sont définis en résolvant (5.8) et (5.12); si $h_j = \infty$, alors $j = J$. En particulier, si $x_{n_j} = 0$ ($y_{n_j} = 0$) alors, $h_j = \infty$, $j = J$ et $x_{n_j+k} = 0$ ($y_{n_j+k} = 0$ ($\forall k \geq 0$)), et l'algorithme prend fin.
- Une fois h_j déterminé, la constante non nulle β_j est donnée par (5.13) et les coefficients $\{\alpha_{s,j}\}_{s=1}^{h_j-1}$ sont obtenus en résolvant le système linéaire triangulaire inférieur (5.14).
- $x_{n_{j+1}}$ et $y_{n_{j+1}}$ sont alors évalués par (5.10).

Les constantes $\gamma_{n,1}$, $\bar{\gamma}_{n,1}$ et les coefficients de récurrence α_m^W , β_m^W dans (5.8) peuvent être choisis arbitrairement. Γ_n et $\gamma_{n,i}$ sont en relation avec $\gamma_{n,1}$ par (5.9) ainsi que $\bar{\Gamma}_n$ et $\bar{\gamma}_{n,i}$ avec $\bar{\gamma}_{n,1}$.

Chapitre 6

Les "breakdowns" de l'algorithme de biorthogonalisation de Lanczos.

Dans ce chapitre, nous reformulerons les résultats principaux des chapitres 1 et 5 à l'aide des relations matricielles obtenues au chapitre 4. Ensuite, nous établirons des liens entre les "breakdowns" de l'algorithme BO et la table de Padé. Enfin, tout comme au chapitre 1, nous formulerons l'algorithme BO pour la résolution des systèmes linéaires, non générique cette fois-ci.

6.1 Formulation matricielle des chapitres 1 et 3.

Soit, la relation obtenue en (4.3)

$$zp(z) = p(z)H,$$

et soit, x_n (y_n) exprimé en fonction du polynôme $P_n(A)$ ($\bar{P}_n(\tilde{A})$) comme en (5.5), formulons le théorème suivant :

Théorème 6.1 *Soit A , une matrice carrée finie. Soit x_0, y_0 , deux vecteurs initiaux orthogonaux. Définissons deux séquences $\{x_n\}_{n=0}^{\infty}$ et $\{y_n\}_{n=0}^{\infty}$ par (5.5) :*

$$\begin{aligned}x_n &= P_n(A)x_0\Gamma_n, \\y_n &= \bar{P}_n(\tilde{A})y_0\bar{\Gamma}_n.\end{aligned}$$

Soit $\{n_j\}_{j=0}^{\infty}$ ($J \leq \infty$), les indices pour lesquels P_n est régulier. Posons

$$\begin{aligned}X &:= [x_0, x_1, x_2, \dots], & Y &:= [y_0, y_1, y_2, \dots], \\ \Gamma &:= \text{diag}[\Gamma_0, \Gamma_1, \Gamma_2, \dots], & \bar{\Gamma} &:= \text{diag}[\bar{\Gamma}_0, \bar{\Gamma}_1, \bar{\Gamma}_2, \dots].\end{aligned}$$

Soit H , la matrice tridiagonale composée de blocs d'Hessenberg supérieurs, définie en (4.2). Définissons H^Γ , $H^{\bar{\Gamma}}$ et D^Γ par :

$$\begin{aligned} H^\Gamma &:= \Gamma^{-1} H \Gamma, \\ H^{\bar{\Gamma}} &:= \bar{\Gamma}^{-1} \bar{H} \bar{\Gamma}, \\ D^\Gamma &:= \bar{\Gamma} D \Gamma. \end{aligned} \tag{6.1}$$

Notons $Y^H X$, la matrice infinie dont l'élément (m, n) est représenté par $\langle y_m, x_n \rangle$. Alors, nous avons

$$\begin{aligned} AX &= XH^\Gamma, \\ A^H Y &= YH^{\bar{\Gamma}} \end{aligned}$$

et

$$Y^H X = D^\Gamma.$$

De plus, notons pour $n \in \mathbb{N}$

$$X_n = [x_0, x_1, \dots, x_n], \quad Y_n = [y_0, y_1, \dots, y_n].$$

Notons H_n^Γ ($H_n^{\bar{\Gamma}}$), la sous-matrice principale d'ordre $(n+1)$ de H^Γ ($H^{\bar{\Gamma}}$) et γ_n^Γ ($\gamma_n^{\bar{\Gamma}}$) l'élément $(n+1, n)$ de cette même matrice infinie H^Γ ($H^{\bar{\Gamma}}$). Enfin, notons e_n^T la dernière ligne de la matrice identité.

Alors,

$$\begin{aligned} AX_n &= X_n H_n^\Gamma + x_{n+1} \gamma_n^\Gamma e_n^T, \\ A^H Y_n &= Y_n H_n^{\bar{\Gamma}} + y_{n+1} \gamma_n^{\bar{\Gamma}} e_n^T. \end{aligned}$$

En sachant que tous les blocs diagonaux finis de D^Γ sont non singuliers et, par conséquent, que x_0, x_1, \dots, x_n , sont linéairement indépendants pour $j < J$, on en vient à la conclusion suivante.

Corollaire 6.2 Si $x_{n_j} = 0$ ($y_{n_j} = 0$) (et donc, $j = J$), alors les colonnes $x_0, x_1, \dots, x_{n_j-1}$ de X_{n_j-1} ($y_0, y_1, \dots, y_{n_j-1}$ de Y_{n_j-1}) sont linéairement indépendantes et engendrent un sous-espace invariant de A (A^H).

A présent, la question est de savoir si la fin de l'algorithme est bien garantie. Le but est de terminer celui-ci avec $x_{n_j} = 0$ ou $y_{n_j} = 0$. Remarquons que le résultat du processus de Lanczos dépend uniquement de A et des deux vecteurs initiaux x_0 et y_0 . L'influence des vecteurs initiaux est double. D'une part, par (5.5), x_0 et y_0 sont les vecteurs sur lesquels $P_n(A)$ et $\bar{P}_n(\bar{A})$ agissent, d'autre part, ils influencent l'ensemble des polynômes orthogonaux formels P_n . Pour en savoir davantage sur l'aboutissement de l'algorithme BO, faisons intervenir les approximations de Padé.

6.2 Les "breakdowns" et la table de Padé.

Au chapitre 2, nous avons travaillé avec les séries fonctionnelles :

$$f(z) := \sum_{k=0}^{\infty} \psi_k z^k$$

et nous avons défini une forme de Padé (p, q) telle que

$$f(z)q(z) - p(z) = \mathcal{O}_+(z^{m+n+1}).$$

Si nous faisons le lien avec l'algorithme BO, nous obtenons la propriété suivante.

Propriété 6.3 *Un "breakdown" de l'algorithme BO se traduira par l'apparition de blocs identiques dans la table de Padé de la fonction f .*

Preuve. L'algorithme BO s'arrêtait lorsque $\delta_\nu = 0$, soit $\beta_\nu = 0$. Or, par (3.8), nous avons

$$P_{\nu+n_j+1}(z) = (z - \alpha_\nu^W)P_{\nu+n_j}(z) - \beta_\nu^W P_{\nu+n_j-1}(z) = (z - \alpha_\nu^W)P_{\nu+n_j}(z).$$

En admettant que Q satisfasse à la même récurrence, nous obtenons

$$Q_{\nu+n_j+1}(z) = (z - \alpha_\nu^W)Q_{\nu+n_j}(z) - \beta_\nu^W Q_{\nu+n_j-1}(z) = (z - \alpha_\nu^W)Q_{\nu+n_j}(z).$$

Par conséquent, les fractions écrites sous forme irréductible :

$$\frac{P_{\nu+n_j}}{Q_{\nu+n_j}}, \quad \frac{P_{\nu+n_j+1}}{Q_{\nu+n_j+1}}$$

sont identiques. □

Théorème 6.4 *Sur la diagonale principale de la table de Padé de la fonction f , l'indice n de la colonne de la première entrée qui est suivie par une entrée identique, est égal à l'indice ν du "breakdown" de l'algorithme BO.*

Il est difficile de tirer des informations de la fonction f relativement à la structure des blocs de sa table. Néanmoins, la détermination de l'ordre exact de f nous permettra de localiser la position du bloc infini, où l'approximant de Padé est égal à f , c'est-à-dire, l'itération où, soit l'algorithme aboutira, soit un "breakdown" surviendra.

Définissons la série

$$F(\xi) := \sum_{k=0}^{\infty} \psi_k \xi^{-k-1}$$

où $\psi_k = \langle y_0, A^k x_0 \rangle$ telle que $f(z) := z^{-1} F(z^{-1})$. Par le théorème de Cauchy-Hadamard, $F(\xi)$ converge au moins pour $|\xi| > \|A\|$. En effet, le théorème de Cauchy-Hadamard nous dit que le rayon de convergence d'une série $\sum_{n=0}^{\infty} a_n x^n$ est

$$\frac{1}{\overline{\lim}_{n \rightarrow \infty} \sqrt[n]{|a_n|}}.$$

La série dont il s'agit ici est

$$F(\xi) = \sum_{k=0}^{\infty} \psi_k \xi^{-k-1}.$$

Nous avons

$$\lim_{k \rightarrow \infty} |\psi_k|^{1/k} \leq \lim_{k \rightarrow \infty} [\|y_0\| \|x_0\| \|A\|^k]^{1/k} = \|A\|.$$

La série converge donc pour $|\xi^{-1}| < 1/\|A\|$, i.e., $|\xi| > \|A\|$.

En insérant ψ_k dans $F(\xi)$, on obtient

$$F(\xi) = \langle y_0, (\xi I - A)^{-1} x_0 \rangle.$$

En effet, par le lemme de Neumann, si $\|T\| < 1$, alors T est inversible et

$$(1 - T)^{-1} = \sum_{n=0}^{\infty} T^n.$$

Comme $F(\xi) := \sum_{k=0}^{\infty} \langle A^k x_0, y_0 \rangle \xi^{-k-1}$ où $\frac{\|A\|}{|\xi|} < 1$, on a

$$I + \xi^{-1} A + \xi^{-2} A^2 + \dots = \xi(\xi I - A)^{-1}$$

de sorte que

$$F(\xi) = \langle y_0, (\xi I - A)^{-1} x_0 \rangle.$$

Nous dirons que $F(\xi)$ est la constante de Schwartz pour le résolvant $(\xi I - A)^{-1}$.

Supposons A diagonalisable, on peut trouver une matrice U non singulière de vecteurs propres telle que

$$AU = UD_{\lambda}$$

où D_{λ} est une matrice diagonale contenant les valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_N$ de A [9]. Ceci donne

$$A^H U^{-H} = U^{-H} D_{\lambda}^H$$

et les colonnes de U^{-H} forment un ensemble de vecteurs propres de A^H . Soit

$$\begin{aligned} U &:= [u_1, u_2, \dots, u_n], \\ U^{-H} &:= V := [v_1, v_2, \dots, v_N] \end{aligned}$$

Représentons x_0 dans la base $\{u_k\}$ et y_0 dans la base $\{v_k\}$:

$$\begin{aligned} x_0 &:= \sum_{k=1}^N u_k \xi_k := U \xi, \\ y_0 &:= \sum_{k=1}^N v_k \eta_k := V \eta \end{aligned}$$

où

$$\begin{aligned} \xi &:= [\xi_1, \xi_2, \dots, \xi_N]^T, \\ \eta &:= [\eta_1, \eta_2, \dots, \eta_N]^T. \end{aligned}$$

Alors, en vue de satisfaire à (5.5), on a

$$\begin{aligned} x_n &= U P_n(D_\lambda) \xi \Gamma_n = \sum_{k=1}^N u_k \Gamma_n P_n(\lambda_k) \xi_k, \\ y_n &= V \bar{P}_n(D_\lambda^H) \eta \bar{\Gamma}_n = \sum_{k=1}^N v_k \bar{\Gamma}_n \overline{P_n(\lambda_k)} \eta_k \end{aligned}$$

et, puisque $V^H U = I$,

$$\begin{aligned} \langle y_m, x_n \rangle &= \bar{\Gamma}_m \eta^H P_m(D_\lambda) P_n(D_\lambda) \xi \bar{\Gamma}_n \\ &= \bar{\Gamma}_m \Gamma_n \sum_{k=1}^N P_m(\lambda_k) P_n(\lambda_k) \xi_k \bar{\eta}_k. \end{aligned}$$

L'hypothèse initiale $\langle y_0, x_0 \rangle \neq 0$ se traduit par

$$\sum_{k=1}^N \xi_k \bar{\eta}_k \neq 0$$

et la condition de "breakdown" sérieux $\langle y_n, x_n \rangle = 0$ peut être exprimée comme

$$\bar{\Gamma}_n \Gamma_n \sum_{k=1}^N P_n^2(\lambda_k) \xi_k \bar{\eta}_k = 0. \quad (6.2)$$

De plus, les moments sont donnés par

$$\psi_k := \langle y_0, A^k x_0 \rangle = \eta^H D_\lambda^k \xi = \sum_{k=1}^N \lambda_k^k \xi_k \bar{\eta}_k.$$

Réécrivons la fonction $F(\xi)$

$$\begin{aligned} F(\xi) &= \langle y_0, (\xi I - A)^{-1} x_0 \rangle \\ &= \eta^H (\xi I - D_\lambda)^{-1} \xi = \sum_{k=1}^N \frac{\xi_k \bar{\eta}_k}{\xi - \lambda_k}. \end{aligned}$$

Soit ν_F , le degré du dénominateur de F tel que ν_F est au plus égal au nombre de termes contenus dans la somme. En vue de satisfaire à

$$\begin{aligned} F(\infty) &= 0, \\ \psi_0 &= \langle y_0, x_0 \rangle \neq 0, \end{aligned}$$

le degré du numérateur de F est $\nu_F - 1$. En effet, nous savons que

$$(\xi I - A)^{-1} = \frac{\text{adj}(\xi I - A)}{\det(\xi I - A)} = \frac{q_{jk}(\xi)}{\chi_A(\xi)} = \frac{\text{polynôme de degré } (N-1)}{\text{polynôme de degré } N}.$$

L'expression de $F(\xi)$ devient

$$\begin{aligned} F(\xi) &= \frac{1}{\chi_A(\xi)} \sum_{j=1}^N \sum_{k=1}^N \underbrace{\tilde{y}_j(\xi) \underbrace{q_{jk}(\xi)}_{\in P_{N-1}} x_k}_{\in P_{N-1}} \\ &= \frac{p_{N-1}(\xi)}{q_N(\xi)}. \end{aligned}$$

De plus, $\lim_{\xi \rightarrow \infty} \xi F(\xi) = \psi_0 \neq 0$.

Nous avons donc bien que le degré du dénominateur est juste d'une unité inférieur au degré du numérateur. Par conséquent, F est de type exact $(\nu_F - 1, \nu_F)$. De plus, si nous supposons que A est non singulière, alors $\xi = 0$ n'est pas un pôle de F , ce qui implique que la fonction f de $z = 1/\xi$, définie par

$$f(\xi^{-1}) = \sum_{k=0}^{\infty} \psi_k \xi^{-k}$$

est de type exact (m, ν_F) avec $m \leq \nu_F$.

Nous savons, par la partie (ii) du théorème de structure en bloc 2.12, que la table de Padé de f possède un bloc infini dont la première colonne est située en $n = \nu_F$ et la première ligne en $m \leq \nu_F$. Ceci implique, par le théorème 2.14, que $\nu_F = n_J$ et $P_{\nu_F} = P_{n_J}$ est le dernier polynôme formel orthogonal pour la fonctionnelle ϕ . Par conséquent, dans la suite, nous pouvons remplacer ν_F par n_J .

Il se peut que (6.2) soit vérifié pour certains $n < n_J$ de telle manière que, dans l'algorithme classique, un "breakdown" sérieux se produise. Nous le dénommerons "curable breakdown". Dans la table de Padé de F , cela correspond à un bloc carré fini et l'algorithme y remédie en sautant au bloc suivant.

Il reste donc à expliquer le "incurable breakdown" où $x_{n_J} \neq 0$, $y_{n_J} \neq 0$ mais par définition de n_J , $\langle x_{n_J}, y_{n_J} \rangle = 0$. Taylor a étudié ce cas en détail [8]. Voici le résultat trouvé.

Théorème 6.8 (Théorème "Mismatch")

Si un "incurable breakdown" survient, i.e. si

$$\langle y_{n_J}, x_{n_J} \rangle = 0, \quad x_{n_J} \neq 0 \text{ et } y_{n_J} \neq 0,$$

le polynôme P_{n_J} est le polynôme caractéristique de la n_J -ième sous-matrice principale $H_{n_J-1}^\Gamma$ de H^Γ et le polynôme minimal d'un sous-espace invariant de A .

6.3 L'algorithme non générique de biorthogonalisation de Lanczos pour la résolution des systèmes linéaires

Dans le cas non générique, le processus BO peut aussi être appliqué à la résolution d'un système linéaire d'équation $Az = b$.

En plus des séquences $\{x_n\}$ et $\{y_n\}$, on génère une séquence $\{z_n\}$ des solutions approximées de telle manière que x_n soit égal au n -ième résidu $(b - Az_n)$. Théoriquement, le but est d'atteindre $x_n = 0$ pour certains n , dans quel cas, z_n sera solution du système. En pratique, la méthode de calcul sera considérée comme une méthode itérative et le but est d'obtenir un résidu suffisamment petit. Il s'agit d'une méthode d'accélération polynomiale puisque le n -ième résidu peut être exprimé comme $x_n = P_n(A)x_0$ où P_n est un polynôme de degré n . Dans une telle méthode, la récurrence pour les résidus peut être écrite sous forme matricielle :

$$AX = XH^\Gamma$$

où H^Γ est une matrice irréductible supérieure d'Hessenberg avec la somme des éléments d'une même colonne égale à 0.

Propriété 6.9 La matrice X respecte la relation $AX = XH^\Gamma$ si et seulement si la somme des éléments dans chaque colonne est nulle.

Preuve. Par analogie avec (1.11), les n premiers pas de la récurrence pour les résidus x_k peuvent toujours être écrits comme

$$AX_n = X_{n+1}H_{n+1,n}^\Gamma$$

où $H_{n+1,n}^\Gamma$ est une matrice $(n+1) \times n$ d'Hessenberg. Une simple transformation donne :

$$X_n = A^{-1}b[1, \dots, 1]H_{n+1,n} - z_{n+1}H_{n+1,n}$$

où $z_{n+1} = [z_0, z_1, \dots, z_n]$. $A^{-1}b[1, \dots, 1]H_{n+1,n} = 0$ si et seulement si la somme des éléments de chaque colonne de $H_{n+1,n}$ est égale à 0. \square

L'algorithme non générique normalisé pour les systèmes linéaires.

Pour résoudre $Az = b$, choisir une approximation initiale z_0 ; poser $x_0 := b - Az_0$; choisir y_0 avec $\langle y_0, x_0 \rangle \neq 0$ et appliquer l'algorithme BO de la section 5.3 avec les formules spéciales suivantes :

- Choisir pour (5.8) les coefficients de récurrence $\{\alpha_m^W\}_{m=0}^\infty$ et $\{\beta_m^W\}_{m=1}^\infty$ ($\beta_0^W := 0$) tels que

$$\gamma_{m+1}^W := \frac{-1}{\alpha_m^W + \beta_m^W \gamma_m^W} \quad (m \in \mathbb{N}) \quad (6.4)$$

où $\gamma_0^W := 1$ et $\gamma_{m+1}^W \neq \infty$.

- Pour $j = 0, 1, \dots$ poser

$$\gamma_{n,1} := \gamma_{n-n_j}^W \quad \text{si } n_j \leq n \leq n_{j+1} - 1, \quad (6.5)$$

$$\gamma_{n_{j+1},1} := \frac{-1}{\alpha_{h_j-1}^W + \beta_{h_j-1}^W \gamma_{n_{j+1}-1,1} + \sum_{i=1}^{h_j} \alpha_{h_j-i,j} \gamma_{n_{j+1}-1,i-1} + \beta_j \gamma_{n_{j+1}-1,h_j-1}} \quad (6.6)$$

- Calculer en plus pour $j = 0, 1, \dots$

$$z_{n+1} := -[x_n + z_n \alpha_{n-n_j}^W] \gamma_{n+1,1} - z_{n-1} \beta_{n-n_j}^W \gamma_{n+1,2} \quad (\text{si } n_j \leq n \leq n_{j+1} - 2) \quad (6.7)$$

$$\begin{aligned} z_{n_{j+1}} &:= -[x_{n_{j+1}-1} + z_{n_{j+1}-1}(\alpha_{h_j-1,j} + \alpha_{h_j-1}^W)] \gamma_{n_{j+1},1} \\ &- z_{n_{j+1}-2}(\alpha_{h_j-2,j} + \beta_{h_j-1}^W) \gamma_{n_{j+1},2} - z_{n_{j+1}-3} \alpha_{h_j-3,j} \gamma_{n_{j+1},3} - \dots \\ &- z_{n_j} \alpha_{0,j} \gamma_{n_{j+1},h_j} - z_{n_j-1} \beta_j \gamma_{n_{j+1},h_j+h_j-1}. \end{aligned} \quad (6.8)$$

Conclusion

Ce mémoire a analysé comment remédier à l'apparition de "breakdowns" dans l'algorithme BO de Lanczos. Pour ce faire, nous sommes passés par le biais de l'étude des polynômes orthogonaux et de la théorie des approximants de Padé.

Quelques résultats importants sont à retenir du chapitre 6 :

- Nous sommes en présence d'un "breakdown" lorsqu'il y a apparition de blocs identiques dans la table de Padé de f .
- L'indice du "breakdown" est donné par l'indice n de la colonne d'un élément de la table de Padé.
- Lorsque le degré de x_0 par rapport à A ne va pas au-delà de l'ordre n_J de la fonction rationnelle f , l'algorithme BO prend fin avec $x_{n_J} = 0$.
- Si un "incurable breakdown" survient, alors le polynôme P_{n_J} est le polynôme caractéristique de la n_J -ième sous-matrice principale et le polynôme minimal d'un sous-espace invariant de A .

Dans ce mémoire, nous avons travaillé avec une fonctionnelle ϕ telle que

$$\phi(z^k) := \psi_k.$$

Il faut savoir qu'il y a moyen de généraliser cette théorie en prenant des fonctionnelles ϕ_l définies $\forall l \in \mathbb{Z}$ par

$$\phi_l(z^k) := \psi_{k+l}.$$

Enfin, il aurait été particulièrement intéressant de procéder à l'implémentation de l'algorithme de Gutknecht pour vérifier la validité de sa théorie sur des expériences numériques. Malheureusement, le temps ne me l'a pas permis.

Bibliographie

- [1] G. GOLUB AND C. VAN LOAN, *Matrix computations*, John Hopkins University Press, Baltimore, 1983.
- [2] W. B. GRAGG, *The Padé table and its relations to certain algorithms of numerical analysis*, SIAM Rev., 14 (1972), pp. 1-62.
- [3] M. H. GUTKNECHT, *Continued fractions associated with the Newton-Padé table*, Interdisciplinary Project Center for Supercomputing, ETH Zurich, ETH-Zentrum, CH-8092 Zurich, 1989.
- [4] M. H. GUTKNECHT, *The unsymmetric Lanczos algorithms and their relations to Padé approximation, continued fractions, and the QD algorithm*, Preliminary Proceedings of the Copper Mountain Conference on Iterative Methods.
- [5] M. H. GUTKNECHT, *A completed theory of the unsymmetric Lanczos process and related algorithms Part I*, SIAM J. Matrix Anal. Appl, 13 (1992), pp.594-639.
- [6] A. S. HOUSEHOLDER, *The Theory of Matrices in Numerical Analysis*, Dover Publications, New-York, 1964.
- [7] H. RUTISHAUSER, *Beiträge zur Kenntnis des Biortogonalisierungs-Algorithmus von Lanczos*, Z. Angew. Math. Phys., 4 (1953), pp. 35-56.
- [8] D. R. TAYLOR, *Analysis of the Look Ahead Lanczos Algorithm*, PhD thesis, Dept. of Mathematics, University of California, Berkeley, 1982.
- [9] J. H. WILKINSON, *The algebraic eigenvalue problem*, Clarendon Press, Oxford, 1965.